MULTILEVEL DIVERSITY CODING
WITH
INDEPENDENT DATA STREAMS

By

HAU Ka Pun

A Thesis

submitted in partial fulfillment of the requirements

for the Degree of Master of Philosophy

Division of Information Engineering

The Chinese University of Hong Kong

June 1995

Acknowledgement

I am pleased to express my deep gratitude to my advisor Dr. Raymond W. H. Yeung for introducing this research topic to me, giving me guidance and encouragement during the whole research period and helping me to polish the drafts of this thesis. In helping me as a beginner in doing research in Information Theory and Diversity Source Coding problems, Dr. Yeung patiently explained some basic concepts to me and teach me some research methods. Also I would like to thank Dr. Zhang Zhen and Dr. Raymond Chan for some very helpful discussions in coding rate regions and extreme points concepts. Also I am grateful to Mr. Ernst Mucke for sending me a compiled programme in convex hull computation which is an important tool in obtaining some results in the thesis.

Besides I would like to thank some of my friends in the Department. They are Sung Chi Wan, Tang Chi Nang, Yip Hoi Man, Poon Hui Chun, Tam Chung Hong and Kwok Ming Shan. They have contributed either directly or indirectly to my work and enriched my life day and night in the Information Engineering Department.

Also I am deeply grateful to Miss Cheung Mun Yee for her concern and her encouragement for me in some difficult times.

I mourn for the death of my friend Sung Chang Hung in the I.E. Department last year. We had pleasant moments of working, talking and playing games together. His death reminded me that there are so many things in one's life that count more than figures and equations.

Abstract

In a Diversity Coding System (DCS), an information source is encoded by a number of encoders, which are then accessed by a number of decoders, each of which is to recover a perfect or distorted version of the information source. In this thesis, we consider the problem of multilevel diversity coding system(MDCS) in which the decoders are divided into different levels and the information source consists of data streams with different levels of importance. Decoders belonging to the same level can recover the same number of data streams.

The coding scheme in an MDCS applying the *principle of superposition* is studied extensively. The code is a special application of the Reed-Soloman code for the distribution of information in different encoders. Each data stream is encoded independently by the encoders and the total coding rates of the encoders are the sum of the coding rates for the individual data streams. This coding scheme is optimal in most cases we have studied but there are a few cases in which it is not optimal. We give examples to illustrate different situations.

In contrast to superposition, another coding scheme in MDCS called *linear combination* are studied and examples are given. We focus our analysis on 3-encoder MDCS's and discover that superposition and linear combination are the *only two* complementary optimal coding schemes in 2-level-3-encoder-MDCS's and superposition and a *hybrid* of superposition and linear combination are the only two complementary optimal coding shemes in 3-level-3-encoder-MDCS's. As a result we can completely characterize the admissible coding rate region for all 2-level and 3-level 3-encoder-MDCS's.

Though the optimality of superposition and linear combination is not mutually conflicting, we discover a class of MDCS's in which the optimality is always not achieved by superposition but by linear combination.

In addition, we study a class of symmetrical MDCS's (SMDCS) which have symmetrical connectivity between encoders and decoders. SMDCS's have special applications in different areas and coding by superposition is optimal in all the cases we have studied. We propose the general admissible coding rate regions and prove the optimality of superposition up to four levels of decoders. We also investigate the relations between different subclasses of SMDCS.

Two equivalent representations of coding rate region of an MDCS induced by superposition, in terms of *rate* constraints and *subrate* constraints respectively, are studied throughout the thesis. They are actually polyhedral sets with special properties. In order to analyze the coding rate regions of MDCS induced by superposition and prove the converse of the coding theorems by means of the rate constraints, we invoke tools and propose some algorithms in convex set analysis, computational geometry and linear programming to analyze the coding rate region.

We hope the work in this thesis can shed some light on further research of the topic.

Contents

Chapter 1

Introduction

1.1 A General Review of MDCS

In a diversity coding system(DCS), an information source is encoded by a number of encoders. There are a number of decoders, each of which is linked to a subset of encoders and is to recover a perfect or distorted version of the information source. The problem is to determine the coding rate region for a particular configuration of a DCS. The following figure is a general configuration of a DCS.

Diversity coding has a wide range of applications. In Roche ,*et al* [11] and Rabin [10] diversity coding is used as a fault tolerance measure in distributed information storage. Instead of storing all the information on a single disk, the information is encoded into a number of pieces , each of which is then stored in a separate disk. In

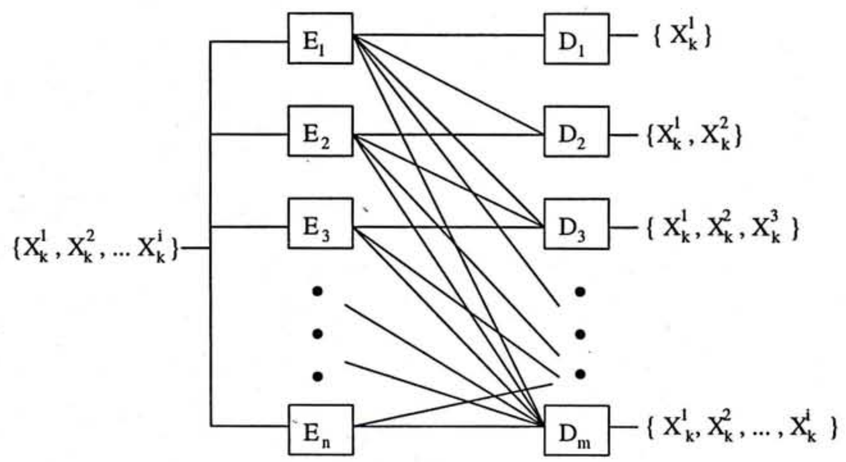


Figure 1.1: A general DCS Configuration

case some of the disks break down, the information can still be recovered from the surviving disks.

In a telecommunication network, information can be transmitted in form of packets which may be lost in case of buffer overflow, misrouting and breakdown of network facilities. If a packet is encoded into several pieces, each taking an independent path to the destination, in case some of the pieces are lost, the information in the packet can still be recovered from the pieces received. Also load balancing, packet delay time and information loss probability can be controlled flexibly by using this scheme. Diversity coding techniques are proposed in Ayanoglu *et al* [1] for general telecommunication network and Lee and Liew [7] for packet transmission in ATM network.

Secret sharing is also an application of diversity coding. In a secret sharing scheme each person(encoder) in a group possesses an encoded version of a secret. The secret is determined only if a certain subgroup of the people (encoders) come together and reveal their knowledge at the same time. Two typical works on secret sharing are Shamir [15] and Karnin [8].

Applications on satellite communication and fault-tolerant database system are proposed by Yeung in [18] and [14] respectively.

Multilevel diversity coding was recently introduced by Roche and Yeung in [12], [18] and [14]. In a multilevel diversity coding system(MDCS), the decoders are classified into different levels. The reconstructions of the source by the decoders of the same level are identical. The information source is encoded in different encoders such that encoders accessing different subset of encoders reconstruct the source with different level of distortion. This kind of system may be implemented for the transmission of information which is decomposable into components of different importance. One example is image and voice (lower frequency component is more important) where a degraded version is still acceptable and the quality can be controlled as the amount of available channels vary in different situations. Roche [12] has defined a class of MDCS called *sequential refinable* which applies similar concept in multilocation information storage. In a sequentially refinable MDCS, the m most important data groups are recovered as some m encoders are accessed. Also it is required in some applications that the reconstructions of the source by decoders within the same level are consistent. An example discussed in Yeung [18] is that the decoders belonging to a certain level form the communication backbone of a distributed control system for a process, and the reconstruction of the source by each encoder is utilized by a control unit which controls a certain part of the process. Then it is very important that the control units act coherently with each other, otherwise the whole process could be impaired. This can be ensured by requiring that all the decoders within the same level deliever the same reconstruction of the source.

In Yeung [18] and Roche *et al* [14], rate distortion approach to MDCS was considered. Reconstructions of the source by decoders within the same levels are subjected to the same distortion criteria. Their results can be specialized to the case in which an information source consists of several independent data streams with different levels of importance. An MDCS can be defined as having different levels of decoders in which the lowest level decoder can recover the most important data stream perfectly in usual Shannon sense. More data streams, in addition to those recovered by lower level decoders, are recovered perfectly by higher level decoders. The decoders of the highest level can recover all the data streams perfectly. We will consider the problem of MDCS with independent data streams based on the problem formulated in Yeung [18] and Roche *et al* [14]. Algebraic coding technique will be used and the optimality of the code is proved information theoretically.

This thesis is mainly composed of four chapters. In Chapter One we give a brief review of multilvel diversity coding system and illustrate the concept of superposition by various examples. In Chapter Two we construct all the possible cases in 2-level and 3-level 3-encoder-MDCS's and examine the optimality of superposition in all of them. We propose optimal coding schemes and characterize the admissible coding rate region for all of them and present some as examples. We focus on MDCS with symmetrical structure in Chapter 3, and investigate different subclasses of SMDCS and how they are related to one another. Some of the tedious proofs are deferred to Appendix A. In Chapter Four we study the coding rate regions of MDCS induced by superposition as polyhedral sets with special properties by invoking mathematical tools from convex set analysis and linear programming and propose short algorithms to solve problems arise in the proof of optimality of superposition. In Appendix B we introduce a class of MDCS's which apply linear combination and prove that superposition is always not optimal in them. Conclusion and suggestion for further research are given in the final chapter of the thesis.

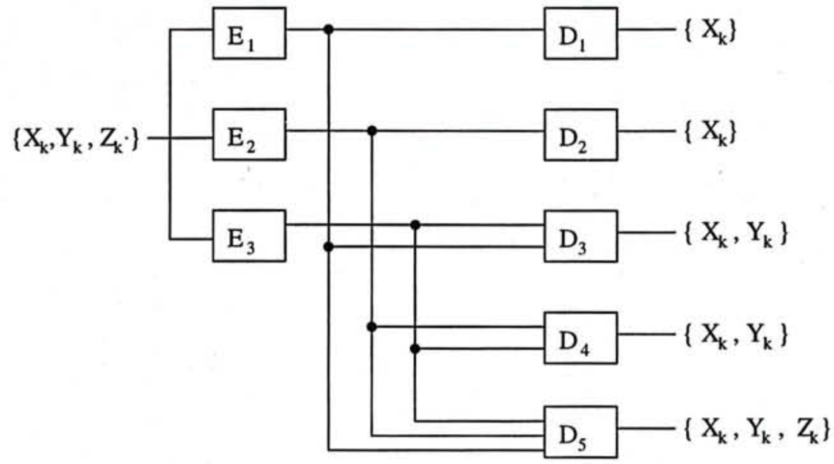


Figure 1.2: An MDCS of 3 encoders and 3 data streams

1.2 MDCS with Independent Data Streams

In this thesis, we consider MDCS's with independent data streams. In such an MDCS, the information source $\{X_k, k = 1, 2, \ldots\}$ is an i.i.d. process. Further, $X_k = (X_k^1, X_k^2, \ldots, X_k^m)$, where $X_k^1, X_k^2, \ldots, X_k^m$ are independent random variables representing independent data streams with descending importance in the source. We use X^i $(1 \le i \le m)$ to denote the generic random variable of $\{X_k^i\}$ and X to denote the generic random variable of $\{X_k\}$. The alphabet set of X^i is denoted by $\mathcal{X}^i$. The source is encoded by a set of encoders indexed by A. The set of decoders are indexed by B. The set of encoders that are accessed by decoder $i \in B$ are indexed by A_i, and this set is called the *fan* of decoder i. The set of decoders that access encoder $i \in A$ are indexed by B_i, and this set is called the *fan* of encoder i. It is assumed that the fan of decoders of different levels are not equal and decoders having the same fan are treated as identical decoders. The *level* of a decoder i is denoted by $l(i)$ and is defined by the number of data streams from $\{X_k^1\}$ up to $\{X_k^m\}$ that the decoder can recover. For example, a decoder i which can recover $\{(X_k^1, \ldots, X_k^j)\}$ belongs to level j, that is $l(i) = j$. The level 1 (lowest level) decoders can recover $\{X_k^1\}$ only. The formulation of the problem here is a special case of the rate distortion approach in [18] and [14] where $\{(X_k^1, X_k^2, ..., X_k^j)\}$, $1 \le j < m$ can be cosidered as a distorted version of the information source.

Example 1.1

Figure 1.2 shows an MDCS with three encoders denoted by $E_i, i \in A = \{1, 2, 3\}$ and 5 decoders denoted by $D_j, j \in B = \{1, 2, 3, 4, 5\}$ and an information source consists of three independent data streams $\{(X_k, Y_k, Z_k)\}$, is encoded in the MDCS. D_1 and D_2 are level 1 decoder which can recover $\{X_k\}$. D_3 and D_4 are level 2 decoder which can recover $\{(X_k, Y_k)\}$. D_5 are level 3 decoder which can recover the whole source $\{(X_k, Y_k, Z_k)\}$. The fans of the encoders and decoders are

$$A_1 = \{1\} \quad A_2 = \{2\} \quad A_3 = \{1, 3\} \quad A_4 = \{2, 3\} \quad A_5 = \{1, 2, 3\}$$

$$B_1 = \{1, 3, 5\} \quad B_2 = \{2, 4, 5\} \quad B_3 = \{2, 3, 5\}$$

$$l(1) = l(2) = 1 \quad l(3) = l(4) = 2 \quad l(5) = 3$$

1.3 Admissible Coding Rate Region

Suppose an i.i.d. source has alpabet set $\mathcal{X}$. A source code of block length n and rate R is a mapping from $\mathcal{X}^n$ into $\{1, 2, ..., M\}$ where M is $\lfloor 2^{nR} \rfloor$. We denote it as an (n, M) code.

In an MDCS with an information source consists of m independent data streams with descending importance, we let S_i as the output random variable of encoder i with alphabet set $\mathcal{S}_i$, An $(n, \{|\mathcal{S}_i|, i \in A\})$ code for each encoder i is defined by

$$f_i : \prod_{k=1}^{m} (\mathcal{X}^k)^n \quad \to \quad \mathcal{S}_i, \quad i \in A$$

and

$$g_j : \prod_{i \in A_j} \mathcal{S}_i \quad \to \quad \prod_{k=1}^{l(j)} (\mathcal{X}^k)^n, \quad j \in B$$

f_i is the encoding function of encoder i and g_j is the decoding function of decoder j accessing encoder i.

Now let the coding rate for encoder i be R_i. A tuple $(R_1, R_2, \ldots, R_{|A|})$ is called *admissible* if there exist, for sufficiently large n, an $(n, \{|\mathcal{S}_i|, i \in A\})$ code such that

$$R_i \geq n^{-1} \log |\mathcal{S}_i| \quad \forall\, i \in A$$

and

$$g_j(f_i((X^k)_1^n, k = 1, \ldots, m), i \in A_j) = g_j(S_i, i \in A_j) = ((X^k)_1^n, k = 1, \ldots, l(j)) \;\forall\, j \in B$$

that is $((X^k)_1^n, k = 1, \ldots, l(j))$ are functions of $(S_i, i \in A_j)$ for all j in B. (The logarithms in this thesis are in base 2 unless otherwise specified.) This is known as the *noiseless coding* of the data streams. An information theoretic description implied by this is,

$$H((X^k)_1^n, k = 1, \ldots, l(j)|S_i, i \in A_j) = 0 \quad \forall\, j \in B \tag{1.1}$$

We will invoke this identity as the consequence of *admissibility* in the prove of converse of a coding scheme.

We define the whole set of admissible tuple as the admissible coding rate region and denote it by $\mathbf{R}$. For the same MDCS, the coding rate region induced by a special kind of feasible coding scheme is denoted by $\mathbf{R}^*$. In general $\mathbf{R}^* \subseteq \mathbf{R}$ because some admissible tuple may be feasible information-theoretically but is not achieved by applying a special kind of coding scheme. The scheme is *optimal* among all the feasible coding schemes if and only if $\mathbf{R}^* = \mathbf{R}$. That is for all the admissible tuple the coding can be done by applying the scheme.

1.4 Distribution of Information in Different Encoders

Suppose in a DCS, the output of an information source is represented in symbol in $GF(q)$, Galios field of with q elements. The source is encoded with block length L by a number of encoders which are the fan of decoder j and are indexed by A_j. Decoder j tries to recover the source from the output of the encoders. Suppose $|A_j| = m$. If an encoder $i, i \in A_j$ uses an (l_i, L) linear block code, then every L-symbol block of the source output, denoted by row L-vector $\mathbf{u}$, is multiplied by a $L \times l_i$ generator matrix $\mathbf{G_i}$ with entries in $GF(q)$ before it is output from encoder i. So the output of encoder i is given by

$$\mathbf{v_i} = \mathbf{u}\mathbf{G_i}$$

By concatenating all $\mathbf{v_i}$ received by decoder j in a single row $\mathbf{v}$, we have

$$\mathbf{v} = [\mathbf{v_1}\ \mathbf{v_2}\ \ldots\ \mathbf{v_m}] = \mathbf{u}[\mathbf{G_1}\ \mathbf{G_2}\ \ldots\ \mathbf{G_m}] = \mathbf{u}\mathbf{G}$$

If we can select L entries in $\mathbf{v}$ and form $\mathbf{v}'$ such that the corresponding L columes in $\mathbf{G}$ forms a nonsigular square $L \times L$ matrix $\mathbf{G}'$, $\mathbf{u}$ can be recovered as

$$\mathbf{u} = \mathbf{v}'(\mathbf{G}')^{-1}$$

To make the scheme feasible, we must find a $\mathbf{G}$ such that some L columns form a non-singular matrix. By applying the results of Reed-Solomon code, we can make $q \geq \sum_{i \in A_j} l_i$ and form a matrix $\mathbf{G}$ with entries in $GF(q)$ as

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_q \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \cdots & \alpha_q^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 & \cdots & \alpha_q^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{L-1} & \alpha_2^{L-1} & \alpha_3^{L-1} & \cdots & \alpha_q^{L-1} \end{bmatrix} \qquad \alpha_k \in GF(q),\ \alpha_i \neq \alpha_j \ \text{ if } i \neq j$$

Any square matrix formed by L distinct columns in $\mathbf{G}$ is the well-known *Vandermonde matrix* and an inverse is guaranteed to exist. In our DCS, the generator matrix $\mathbf{G_i}$ of encoder i can be formed by concatenate l_i distinct columns in $\mathbf{G}$.

Fault-tolerance is implied by the above scheme. We can make $\sum_{i \in A_j} l_i$ larger than L as we wish and if some of the encoders break down or some of the communication channels linking the encoders and decoder are blocked, we can still recover the information vector $\mathbf{u}$ as long as we can obtain L distinct symbols from the surviving encoders. In fact in most literatures, similar methods are applied to construct divesity codes.([10], [1], [7] and [18]).

Now suppose the rate of the source is R symbols in $GF(q)$ per unit time. For a encoder i with an $l_i \times L$ generator matrix, the coding rate of the encoder i contributing to encode the source is given by

$$r_i = R(\log q)\frac{l_i}{L} \text{ (bits per second)}$$

If the source has entropy rate H, by Shannon's Source Coding Theorem [16], the decoder can reconstruct the source perfectly if and only if

$$\sum_{i \in A_j} r_i \geq H \tag{1.2}$$

1.5 Multilevel Diversity Coding by Superposition

Generally in an MDCS, an information source consists of more than one data streams. For example, an information source $\{(X_k, Y_k)\}$, where $\{X_k\}$ and $\{Y_k\}$ are independent data streams, is to be encoded by a set of encoders in an MDCS. The fan of the decoder that is to recover $\{X_k\}$ is A_x and the fan of the decoder that is to recover $\{Y_k\}$ is A_y. Now if $\{X_k\}$ and $\{Y_k\}$ are independent, one may expect the optimality can be achieved by coding $\{X_k\}$ and $\{Y_k\}$ separately. The argument is that the coding rate contributing to the coding of $\{X_k\}$ does not contribute to the coding of $\{Y_k\}$. That is, the coding rate R_j of an encoder j in which the two streams are encoded is given by

$$R_j = r_j^x + r_j^y$$

Note that for each data scheme, we apply the coding scheme described in the last section. Now in this context R_j is refered to as the *rate* of encoder j and r_j^x, r_j^y are refered to as *subrate* of encoder j contributing to encode $\{X_k\}$ and $\{Y_k\}$ respectively. In such a coding scheme, for perfect recovery of $\{X_k\}$ and $\{Y_k\}$, by (1.2) we must have

$$\sum_{i \in A_x} r_i^x \geq H(X)$$
$$\sum_{i \in A_y} r_i^y \geq H(Y)$$

This coding scheme is first proposed in Yeung [18] as a general coding scheme in an MDCS. This coding scheme does not involve any arithmetic manipulation between different data streams and is conceptually easy to implement. It is always one of the feasible coding schemes. We will see that under some situation it is also an optimal scheme. If this is the case, we say that *the principle of superposition applies*, or *superposition is optimal.* We will look into the details of this problem.

In an MDCS with $|A|$ encoders and an information source consists of m independent data streams represented by $\{(X_k^1, X_k^2, \ldots, X_k^m)\}$, for any rate tuple $(R_1, R_2, \ldots, R_{|A|})$ in the coding rate region induced by superposition, we have

$$R_i = r_i^1 + r_i^2 + \ldots + r_i^m \quad \text{where } r_i^j \geq 0 \text{ for } 1 \leq i \leq |A| \text{ and } 1 \leq j \leq m.$$

Suppose those decoders which recover stream $\{X_k^i\}$ are indexed by B_i, fan of an decoder j in B_i is A_{ij}. We must have

$$\sum_{k \in A_{ij}} r_k^i \geq H(X^i) \quad \text{for } 1 \leq i \leq m, \ \forall\, j \in B_i$$

Example 1.1 continued

Let us reconsider the MDCS in Example 1.1.

The coding rate region induced by superposition is given by

$\{(R_1, R_2, R_3) : \ r_i^x, r_i^y, r_i^z \geq 0, \text{ for } i = 1, 2, 3 \text{ and}$

$$\begin{aligned}
R_i &= r_i^x + r_i^y + r_i^z \text{ for } i = 1,2,3 && (1.3)\\
r_1^x &\geq H(X) && (1.4)\\
r_2^x &\geq H(X) && (1.5)\\
r_1^y + r_3^y &\geq H(Y) && (1.6)\\
r_2^y + r_3^y &\geq H(Y) && (1.7)\\
r_1^z + r_2^z + r_3^z &\geq H(Z) \qquad \} && (1.8)
\end{aligned}$$

The following inequalities

$$\begin{aligned}
r_1^x + r_3^x &\geq H(X)\\
r_2^x + r_3^x &\geq H(X)\\
r_1^x + r_2^x + r_3^x &\geq H(X)\\
r_1^y + r_2^y + r_3^y &\geq H(Y)
\end{aligned}$$

are implied by $(1.4), (1.5)$and(1.6) respectively and so are not shown in the rate region description for a compact description.

We will see shortly that an equivalent description of the same coding rate region is given by

$\{(R_1, R_2, R_3) : \ R_i \geq 0 \ \text{ for } i = 1, 2, 3 \ \text{ and}$

$$\begin{aligned}
R_1 &\geq H(X) && (1.9)\\
R_2 &\geq H(X) && (1.10)\\
R_1 + R_3 &\geq H(X) + H(Y) && (1.11)
\end{aligned}$$

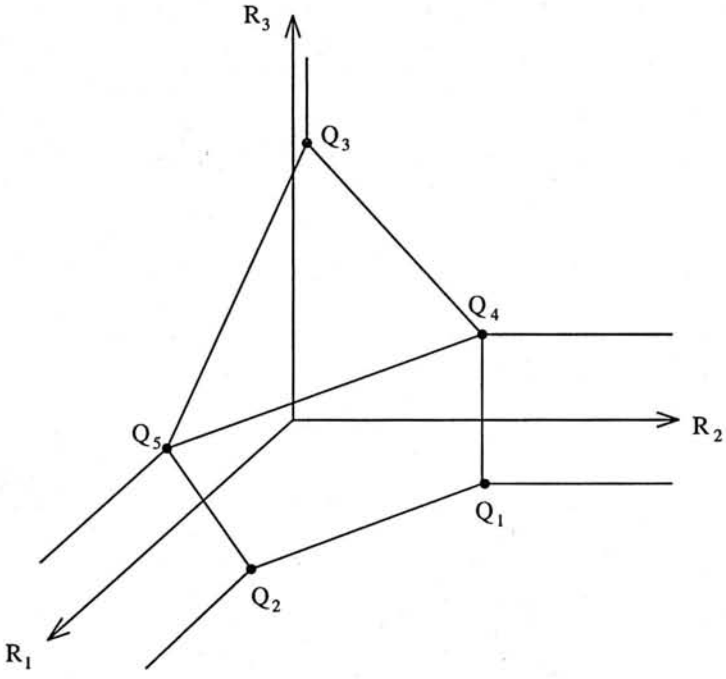


Figure 1.3: Coding Rate Region of the MDCS in Example 1.1

$$
\left.\begin{array}{rcl}
R_2 + R_3 & \geq & H(X) + H(Y) \\
R_1 + R_2 + R_3 & \geq & 2H(X) + H(Y) + H(Z) \\
R_1 + R_2 + 2R_3 & \geq & 2H(X) + 2H(Y) + H(Z)
\end{array}\right\}
\begin{array}{r}(1.12)\\(1.13)\\(1.14)\end{array}
$$

This set is shown in the Figure 1.3.

We denote the first set of inequalities as *subrate constraints* and the coding rate region induced by superposition described by them as $\mathbf{r_{sp}}$. The second set of inequalities is denoted as the *rate constraints* and the corresponding region as $\mathbf{R_{sp}}$. Once the MDCS is defined, the set of subrate constraints are immediately determined. But establishing an equivalent set of rate constraints is a nontrivial process. We will look into this problem in chapter 4 with insights from convex set analysis. But first let us see how we can prove that the two descriptions of coding rate regions above are equivalent. We will first prove $\mathbf{r_{sp}} \subseteq \mathbf{R_{sp}}$ by showing $(1.4)-(1.8)$ imply $(1.9)-(1.14)$ given that $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3$ such that (1.3) is satisfied.

$$
\begin{array}{rcl}
(1.4) & \Rightarrow & (1.9) \\
(1.5) & \Rightarrow & (1.10) \\
(1.4) + (1.6) & \Rightarrow & (1.11) \\
(1.5) + (1.7) & \Rightarrow & (1.12)
\end{array}
$$

$$\begin{aligned}(1.4)+(1.5)+(1.6)+(1.8) &\Rightarrow (1.13)\\ (1.4)+(1.5)+(1.6)+(1.7)+(1.8) &\Rightarrow (1.14)\end{aligned}$$

To prove $\mathbf{r_{sp}} \subseteq \mathbf{R_{sp}}$ we invoke Theorem 8 in Chapter Four. It is seen that any triple in $\mathbf{R_{sp}}$ is also in $\mathbf{r_{sp}}$ if the all the extreme points of $\mathbf{R_{sp}}$ is in $\mathbf{r_{sp}}$. Using the algorithm introduced in Chapter Four we enumerate all the extreme points of $\mathbf{R_{sp}}$, denoted each as Q_i and show that they are in fact in $\mathbf{r_{sp}}$ by expressing each as a sum of three triples satisfying the subrate constraints in $\mathbf{r_{s}p}$.

$$\begin{aligned}
Q_1 &= (H(X)+H(Y), H(X)+H(Y)+H(Z), 0)\\
&= (H(X), H(X), 0) + (H(Y), H(Y), 0) + (0, H(Z), 0)\\
Q_2 &= (H(X)+H(Y)+H(Z), H(X)+H(Y), 0)\\
&= (H(X), H(X), 0) + (H(Y), H(Y), 0) + (H(Z), 0, 0)\\
Q_3 &= (H(X), H(X), H(Y)+H(Z))\\
&= (H(X), H(X), 0) + (0, 0, H(Y)) + (0, 0, H(Z))\\
Q_4 &= (H(X), H(X)+H(Z), H(Y))\\
&= (H(X), H(X), 0) + (0, 0, H(Y)) + (0, H(Z), 0)\\
Q_5 &= (H(X)+H(Z), H(X), H(Y))\\
&= (H(X), H(X), 0) + (0, 0, H(Y)) + (H(Z), 0, 0)
\end{aligned}$$

So $\mathbf{r_{sp}} = \mathbf{R_{sp}}$

The significance of the rate constraints is that they are necessary in proving the optimality of superposition. This is elaborated in the next section.

1.6 Optimality of Superposition

Obviously $\mathbf{R_{sp}} \subseteq \mathbf{R}$ since superposition is a feasible coding scheme. Now the question is how we can prove that superposition is an optimal scheme in some situation. If we can also establish $\mathbf{R} \subseteq \mathbf{R_{sp}}$ (the converse of the coding theorem) which is to show that every admissible tuple is also inside $\mathbf{R_{sp}}$ then $\mathbf{R_{sp}} = \mathbf{R}$ and the optimality of superposition is proved. We are unable to establish the converse part by means of the coding rate region $\mathbf{r_{sp}}$ described in terms of subrate constraints. Instead we make use of the coding rate region $\mathbf{R_{sp}}$ described in terms of the rate constraints to do so.

Example 1.1 continued

Now we use $\mathbf{R_{sp}}$ of Example 1.1 to prove the converse.
For any admissible triple (R_1, R_2, R_3),

$$\begin{aligned}
R_1 &\geq n^{-1}H(S_1) && (1.15)\\
&= n^{-1}(H(S_1) + H(X_1^n|S_1)) && (1.16)\\
&= n^{-1}H(S_1, X_1^n)\\
&= n^{-1}H(S_1|X_1^n) + n^{-1}H(X_1^n)\\
&\geq H(X)\\
\text{similarly, } R_2 &\geq H(X)\\
R_1 + R_3 &\geq n^{-1}(H(S_1) + H(S_3)) && (1.17)\\
&= n^{-1}(H(S_1|X_1^n) + H(X_1^n) + H(S_3)) && (1.18)\\
&\geq n^{-1}(H(S_1|X_1^n) + H(S_3|X_1^n)) + H(X) && (1.19)\\
&\geq n^{-1}(H(S_1, S_3|X_1^n)) + H(X)\\
&= n^{-1}(H(S_1, S_3, Y_1^n|X_1^n)) + H(X) && (1.20)\\
&= n^{-1}H(S_1, S_3|X_1^n, Y_1^n) + H(Y|X) + H(X)\\
&\geq H(Y) + H(X)\\
\text{similarly, } R_2 + R_3 &\geq H(X) + H(Y)\\
R_1 + R_2 + R_3 &\geq n^{-1}(H(S_1) + H(S_2) + H(S_3)) && (1.21)\\
&= n^{-1}(H(S_1|X_1^n) + H(S_2|X_1^n) + H(S_3)) + 2H(X) && (1.22)\\
&\geq n^{-1}(H(S_1|X_1^n) + H(S_2|X_1^n) + H(S_3|X_1^n)) + 2H(X) && (1.23)\\
&\geq n^{-1}H(S_1, S_2, S_3, Y_1^n, Z_1^n|X_1^n) + 2H(X) && (1.24)\\
&= n^{-1}H(S_1, S_2, S_3|X_1^n, Y_1^n, Z_1^n) + 2H(X) + H(Y) + H(Z)\\
&\geq 2H(X) + H(Y) + H(Z)\\
R_1 + R_2 + 2R_3 &\geq n^{-1}(H(S_1) + H(S_2) + 2H(S_3)) && (1.25)\\
&= n^{-1}(H(S_1|X_1^n) + H(S_2|X_1^n) + 2H(S_3)) + 2H(X) && (1.26)\\
&\geq n^{-1}(H(S_1|X_1^n) + H(S_2|X_1^n) + 2H(S_3|X_1^n)) + 2H(X) && (1.27)\\
&\geq n^{-1}(H(S_1, S_3, Y_1^n|X_1^n) + H(S_2, S_3, Y_1^n|X_1^n)) + 2H(X) && (1.28)\\
&= n^{-1}(H(S_1, S_3|X_1^n, Y_1^n) + H(S_2, S_3|X_1^n, Y_1^n)) + 2H(X) + 2H(Y)\\
&\geq n^{-1}(H(S_1, S_2, S_3, Z_1^n|X_1^n, Y_1^n) + 2H(X) + 2H(Y) && (1.29)
\end{aligned}$$

$$
\begin{aligned}
&\geq n^{-1}(H(S_1,S_2,S_3|X_1^n,Y_1^n,Z_1^n)+2H(X)+2H(Y)+H(Z) \\
&\geq 2H(X)+2H(Y)+H(Z)
\end{aligned}
$$

The validity of the above equalities and inequalities are explained in the following.(In the rest of the thesis these explanations will be much simplified)

1. (1.15), (1.17), (1.21) and (1.25) hold because by admissibility condition, $R_i \geq n^{-1}\log|\mathcal{S}_i|$ which is an upper bound of $n^{-1}H(S_i)$.

2. (1.16), (1.18), (1.20), (1.22), (1.24), (1.26), (1.28) and (1.29) hold because by the admissibility condition, (X_1^n) is a S_1 and is also a function of S_2, (Y_1^n) is a function of (S_1,S_2) and is also a function of (S_2,S_3) and (Z_1^n) is a function of (S_1,S_2,S_3).

3. (1.19), (1.23) and (1.27) holds because conditioning reduce entropy.

Since any admissible tuple satisfies the constraints describing $\mathbf{R_{sp}}$, it is inside $\mathbf{R_{sp}}$. So $\mathbf{R} \subseteq \mathbf{R_{sp}}$. The converse of the coding theorem is true. This proves that $\mathbf{R} = \mathbf{R_{sp}}$. So superposition is optimal in this case of MDCS.

But in general superposition is a feasible, but not necessarily optimal coding scheme. In the following we show 3 examples of MDCS which illustrate the subtlety of MDCS problems. We can prove superposition is not optimal in the first two of them by giving a better coding scheme instead of superposition. The third example of MDSC is a slight modification of the second but superposition is found to be optimal in that case. The three MDCS's' are shown in Figure 1.4.

Example 1.2

We now define the coding rate region $\mathbf{r_{sp}}$ for the MDSC as
$\{(R_1,R_2,R_3):\ R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i=1,2,3$ and

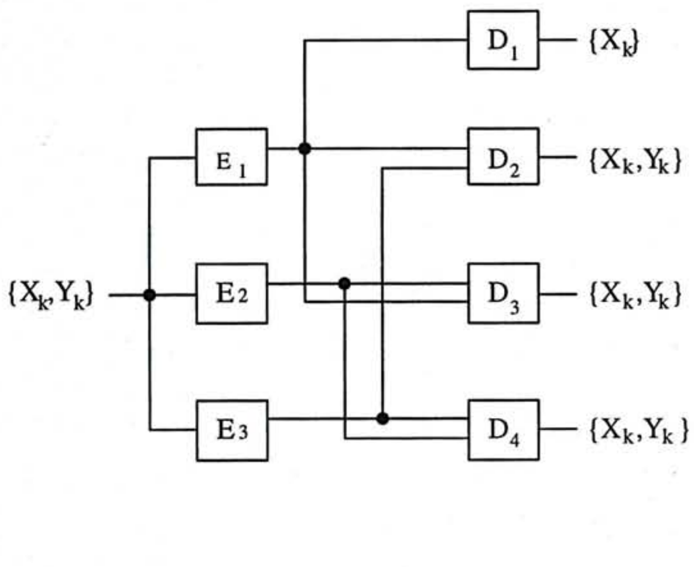
$$r_1^x \geq H(X) \tag{1.30}$$

$$r_2^x + r_3^x \geq H(X) \tag{1.31}$$

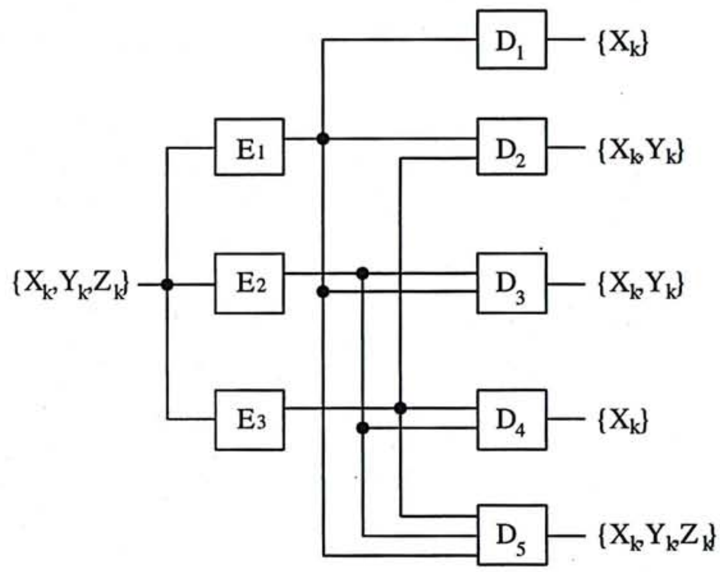
$$r_1^y + r_2^y \geq H(Y) \tag{1.32}$$

$$r_2^y + r_3^y \geq H(Y) \tag{1.33}$$

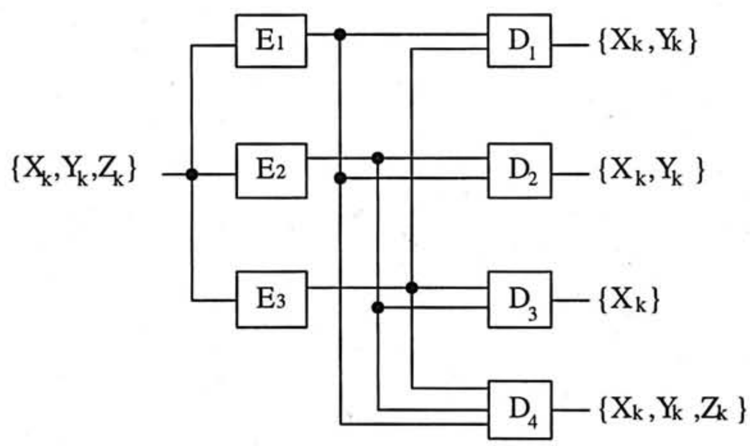
$$r_1^y + r_3^y \geq H(Y) \quad \} \tag{1.34}$$



MDCS of Example 1.2



MDCS of Example 1.3



MDCS of Example 1.4

Figure 1.4: The Configuration of three MDCS's

An equivalent coding rate region $\mathbf{R_{sp}}$ is
$\{(R_1, R_2, R_3):\ R_i \geq 0$ for $i = 1,2,3$ and

$$\left.\begin{aligned} R_1 &\geq H(X) && (1.35)\\ R_1+R_2 &\geq H(X)+H(Y) && (1.36)\\ R_2+R_3 &\geq H(X)+H(Y) && (1.37)\\ R_1+R_3 &\geq H(X)+H(Y) && (1.38)\\ 2R_1+R_2+R_3 &\geq 3H(X)+2H(Y) \quad \} && (1.39)\end{aligned}\right.$$

We let $\{X_k\}, \{Y_k\}$ be independent bits with

$$p[X_k = 0] = p[X_k = 1] = p[Y_k = 0] = p[Y_k = 1] = 0.5$$

So $H(X) = H(Y) = 1$. The triple $(1,1,1)$ is admissible by coding $\{X_k\}, \{Y_k\}$ in encoder 1, 2 and 3 as $\{X_k\}, \{Y_k\}, \{X_k \oplus Y_k\}$ respectively. But it is outside $\mathbf{R_{sp}}$ since it does not satisfy (1.39).

Example 1.3

$\mathbf{r_{sp}}$ is given by
$\{(R_1, R_2, R_3):\ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} r_1^x &\geq H(X) && (1.40)\\ r_2^x + r_3^x &\geq H(X) && (1.41)\\ r_1^y + r_2^y &\geq H(Y) && (1.42)\\ r_2^y + r_3^y &\geq H(Y) && (1.43)\\ r_1^z + r_2^z + r_3^z &\geq H(Z) \quad \} && (1.44)\end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by
$\{(R_1, R_2, R_3):\ R_i \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} R_1 &\geq H(X) && (1.45)\\ R_2+R_3 &\geq H(X) && (1.46)\\ R_1+R_3 &\geq H(X)+H(Y) && (1.47)\\ R_2+R_3 &\geq H(X)+H(Y) && (1.48)\\ R_1+R_2+R_3 &\geq 2H(X)+H(Y)+H(Z) && (1.49)\\ 2R_1+R_2+R_3 &\geq 3H(X)+2H(Y)+H(Z) \quad \} && (1.50)\end{aligned}$$

Again we let $\{X_k\}, \{Y_k\}$ be independent bits with

$$p[X_k = 0] = p[X_k = 1] = p[Y_k = 0] = p[Y_k = 1] = 0.5$$

So $H(X) = H(Y) = 1$. Also assume $H(Z) = 1$. By coding $\{X_k\}, \{Y_k\}$ in the three encoders as $\{X_k\}, \{Y_k\}$ and $\{X_k \oplus Y_k\}$ and rate of coding $\{Z_k\}$ is added onto each encoder i as r_i^z such that $r_1^z + r_2^z + r_3^z \geq 1$. So the triple $(1, 1, 2)$ is admissible. But it is outside $\mathbf{R_{sp}}$ since it does not satisfy (1.50).

Example 1.4

$\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, R_3): \ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_2^x + r_3^x &\geq H(X) && (1.51)\\ r_1^x + r_2^x &\geq H(X) && (1.52)\\ r_1^x + r_3^x &\geq H(X) && (1.53)\\ r_1^y + r_3^y &\geq H(Y) && (1.54)\\ r_2^y + r_3^y &\geq H(Y) && (1.55)\\ r_1^z + r_2^z + r_3^z &\geq H(Z) \qquad \} && (1.56) \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by

$\{(R_1, R_2, R_3): \ R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_2 + R_3 &\geq H(X) && (1.57)\\ R_1 + R_2 &\geq H(X) + H(Y) && (1.58)\\ R_1 + R_3 &\geq H(X) + H(Y) && (1.59)\\ R_1 + R_2 + R_3 &\geq \frac{3}{2}H(X) + H(Y) + H(Z) && (1.60)\\ 2R_1 + R_2 + R_3 &\geq 2H(X) + 2H(Y) + H(Z) && (1.61)\\ R_1 + 2R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) && (1.62)\\ R_1 + R_2 + 2R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \} && (1.63) \end{aligned}$$

The MDCS in Example 1.4 is a slight modification of the one in Example 1.3 by removing D_1 which recovers $\{X_k\}$, but it turns out that superposition is optimal in this example. It can be easily shown that any admissible triple is inside $\mathbf{R_{sp}}$ by proving the converse. Moreover it is still feasible by coding $\{X_k\}, \{Y_k\}$ and $\{Z_k\}$ the same as the one in Example 1.3 and the rate tuple (1,1,2) is inside the coding rate region $\mathbf{R_{sp}}$.

1.7 Different MDCS coding schemes

Here we briefly introduce different feasible MDCS coding schemes while superposition is one of them.

If we consider all data streams as bit streams, then the linear combination of different streams is just the exclusive-or $\oplus$ operation, or the addition operation in GF(2), on the bits. As we can see from the examples in the last section, another feasible coding scheme in MDCS is linearly combining different data streams and output as a single stream. In the rest of the thesis, this coding scheme is generally termed as *linear combination* in contrast to *superposition.*

For a bit stream $\{X_k\}$ with entropy rate $H(X)$, we can regard it as a bit stream compressed to a maximum extend in a lossless compression and the output is $H(X)$ bits per second. As for a maximally compressed bit streams, the probability of 0 and 1 are equally likely (otherwise it can be further compressed). If it is linearly combined with another (also assume maximally compressed) independent bit stream $\{Y_k\}$ with entropy rate $H(Y)$, the the first $H(X)$ bits of the output is the bit-wise exclusive-or of the bits of $\{X_k\}$ and first $H(X)$ bits of $\{Y_k\}$ and the remaining $H(Y) - H(X)$ bits of the output are just the last $H(Y) - H(X)$ bits of $\{Y_k\}$. For the exclusive-or operation between two bits in which 0 and 1 are equally likely, it is easy to see that 0 and 1 are also equally likely in the resulting bit. So the resulting bit stream $\{X_k \oplus Y_k\}$ cannot be further compressed and the entropy rate is just $H(Y)$, the maximum of the entropy rates of the two streams. The same argument applies to the rest of the thesis.

In Roche [12] a class of MDCS which is called *sequential refinable* was introduced. The problem originated in multilocation information storage system. Suppose in a network with m users and n disks each disk with capacity C bits, the total information stored in the network is $k_{max}C$ bits and is represented by the vector $[W_1\ W_2\ \ldots\ W_{k_{max}}]$ where W_i is the variable representing the ith C bits of information. Each user which can access k_i disks ($k_i \leq k_{max}$) can reliably recover the first k_iC bits of information $[W_1, W_2, \ldots, W_{k_i}]$.

The coding of the sequential refinable MDCS is done by storing *linear combinations*

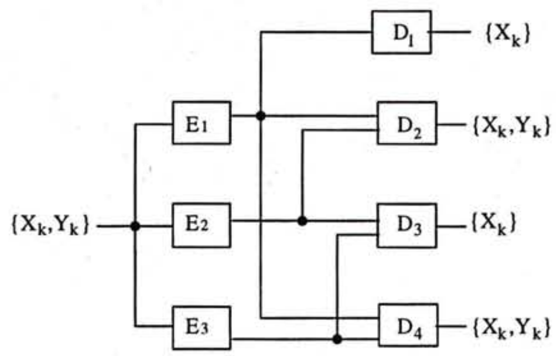


Figure 1.5: MDCS where linear combination applies

of variables W_i (exclusive-or of W_i) in some of the disks instead of storing the variables themselves. Example 1.2 above just shows a typical example of sequential refinable MDCS. Since the number of information bits one can recovered is equal to the number of bits of information stored in the accessible disks. It is an *optimal* coding scheme. However, the configuration of MDCS must satisfy certain constraints before the coding scheme can be implemented [12]. So the scheme is optimal but not necessarily feasible in an MDCS. (In contrast to superposition which is always feasible but not necessarily optimal).

Note the subtlety is that sequentially refinable MDCS's are just one special class of MDCS which apply *linear combination.* In figure 1.5, we give a simple example of MDCS which applies linear combination but does not belong to the class of sequentially refinable MDCS. In this example, the information source $\{(X_k, Y_k)\}$ can be coded in encoder 1, 2 and 3 as $\{X_k\}$, $\{Y_k\}$ and $\{X_k \oplus Y_k\}$ respectively. It is not sequentially refinable since decoder 3 connected to encoder 2 and encoder 3 recovers $\{X_k\}$ only. This example and Example 1.2 are members of a general class of MDCS for which optimality can be acheived by linear combination and *superposition is always not optimal.* This class of MDCS will be discussed in Appendix B.

Note the subtlety is that the optimality of superposition and linear combination are not mutually conflicting. The same optimal rate tuple can be achievable by applying superposition or by linear combination. In Roche [12], one MDCS was used as example to illustrate the concept of coding in a sequentially refinable MDCS. The MDCS is shown in Figure 1.6. The name of the variables are modified to fit our context.

It was concluded in [12] that the three independent bit streams $\{X_k, Y_k, Z_k\}$, each

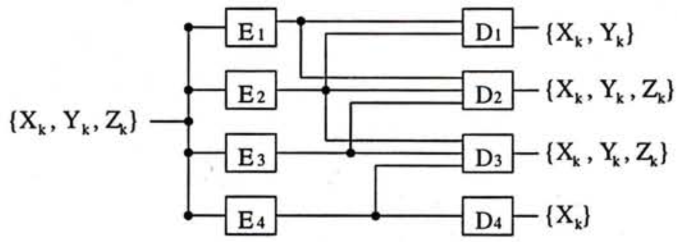


Figure 1.6: MDCS where superposition and linear combination is both optimal

with entropy rate C bits, can be coded as $\{X_k \oplus Y_k\}, \{Y_k\}, \{Z_k\}, \{X_k\}$ in encoder 1, 2, 3 and 4 respectively and the rate is (C, C, C, C). While it is a feasible coding scheme, we discover that it is also feasible by simply putting $\{X_k\}, \{Y_k\}, \{Z_k\}, \{X_k\}$ in the 4 encoders and also achieve a rate of (C, C, C, C). So the rate achievable by superposition and linear linear combination is the same and coding by linear combination seems unnecessary.

The coding scheme of the MDCS in Example 1.3 above combines both superposition and linear combination of streams while coding by superposition alone is not optimal and applying linear combination alone is infeasible. In Example 1.4, coding can be done by superposition alone or by combining superposition and linear combination. The rate tuples induced by both are inside $\mathbf{R_{sp}}$. We will see that all such hybrid coding schemes plays an important role in 3-level-3-encoder MDCS's. In fact superpostion, linear combination and hybrid of the two are three complementary optimal coding schemes in all cases of 2-level and 3-level 3-encoder-MDCS's.

Chapter 2

MDCS's with Three Encoders

We have seen that *superposition* in general is a feasible coding scheme in an MDCS. In this chapter we are going to analyze what coding scheme is optimal in an MDCS. We will focus on MDCS's with three encoders and the decoders belongs to two or three different levels. (In the following, a 2-level-3-encoder MDCS means an MDCS with there encoders, each belonging to one of two different levels.) We let the information source $\{(X_k, Y_k)\}$ to be encoded in a 2-level MDCS where $\{X_k\}$ and $\{Y_k\}$ are independent data streams. A level-1 decoder recovers $\{X_k\}$ perfectly and a level-2 decoder recovers $\{(X_k, Y_k)\}$ perfectly. In the 3-level MDCS, $\{(X_k, Y_k, Z_k)\}$ are encoded and a level-1 decoder recovers $\{X_k\}$, a level-2 decoder recovers $\{(X_k, Y_k)\}$ and a level-3 decoder recovers $\{(X_k, Y_k, Z_k)\}$ perfectly. Also the three data streams are independent.

Since there are insufficient insight at the moment on the general situation of the different coding schemes or how the admissible coding rate regions are like, we start with an empirical study on all the possible cases of 2-level and 3-level 3-encoder-MDCS's. After enumerating all the possible configurations, we examine whether superposition is optimal in them. If superposition is optimal, then the coding rate region $\mathbf{R_{sp}}$ is the general admissible region. If it is not, we will propose an alternative coding scheme which can achieve the optimal coding rate. By finding suitable rate constraints and their corresponding extreme rate tuples, we try to characterize the admissible coding rate for all those cases. It turns out that we can characterize the general admissible coding rate regions for all 2-level and 3-level 3-encoder MDCS's. Note that we can generalize our results up to 5-level-3-encoder MDCS's but we do

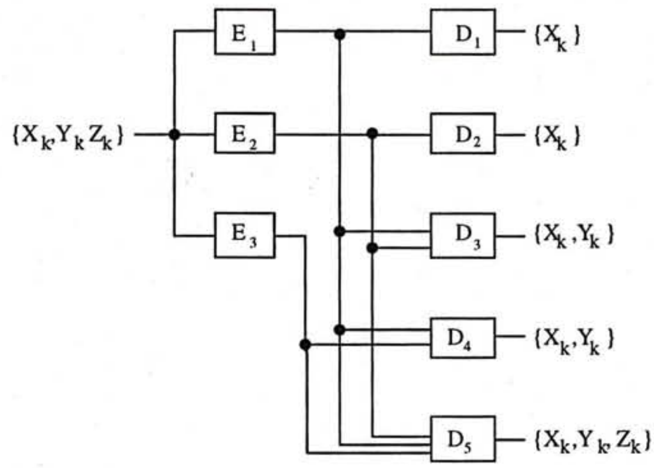


Figure 2.1: MDCS of Example 2.1

not consider thoses cases at the moment.

The following notation will be used in the whole chapter, the three encoders of an MDCS are indexed by 1, 2 and 3 respectively. The fan of a decoder is grouped in small blackets (). Each index is separated by a comma. The fans of different decoders within the same level are grouped in large blackets {} while the fan of each decoder is separated by a comma. This notation uniquely determines the configuration of a 3-encoder MDCS.

Example 2.1

The MDCS denoted as

$$\begin{array}{ccc} level\ 1 & level\ 2 & level\ 3 \\ \{(1),(2)\} & \{(1,2),(1,3)\} & \{(1,2,3)\} \end{array}$$

has a unique configuration as shown in Figure 2.1.

Now we investigate the 2-level-3-encoder-MDCS's first.

2.1 2-level-3-encoder MDCS

The different configurations of 2-level-3-encoder MDCS are constructed according the following procedure.

First the fans of level-1 decoders are identified. There are a total of 7 different structures that are feasible for an MDCS with decoders belonging to more than one level. :

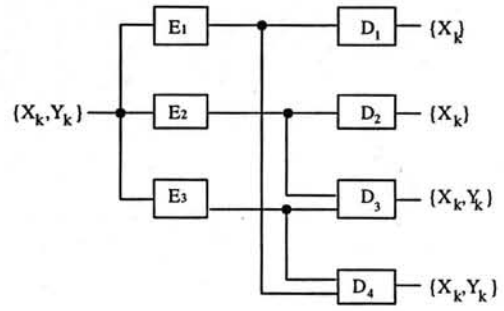
$$\{(1)\}, \{(1),(2)\}, \{(1),(2),(3)\}, \{(1,2)\}, \{(1),(2,3)\}, \{(1,2),(2,3)\}, \{(1,2),(1,3),(2,3)\}$$

The structures symmetrical to them are treated as identical. For example, $\{(1,2),3\}$ is symmetrical to $\{(1,(2,3)\}$ so it is not considered. Then the all the possible fans of level-2 decoders are identified. They can be neither identical to the fan of any level-1 decoder nor a subset of it.

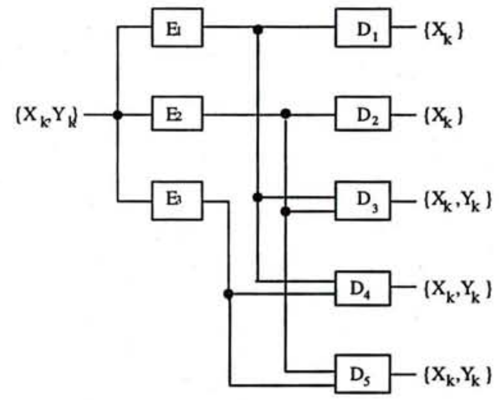
In the following we list all the possible configuration of a 2-level-3-encoder MDCS's.

The (*) indicates that superposition is not optimal in the MDCS.

	level 1	*level* 2
1.	$\{(1)\}$	$\{(2)\}$
2.	$\{(1)\}$	$\{(2),(3)\}$
3.	$\{(1)\}$	$\{(1,2)\}$
4.	$\{(1)\}$	$\{(2,3)\}$
5.	$\{(1)\}$	$\{(1,2),(3)\}$
6.	$\{(1)\}$	$\{(1,2),(2,3)\}$
7.	$\{(1)\}$	$\{(1,2),(1,3)\}$
8.*	$\{(1)\}$	$\{(1,2),(1,3),(2,3)\}$
9.	$\{(1)\}$	$\{(1,2,3)\}$
10.	$\{(1),(2)\}$	$\{(3)\}$
11.	$\{(1),(2)\}$	$\{(1,2)\}$
12.	$\{(1),(2)\}$	$\{(2,3)\}$
13.	$\{(1),(2)\}$	$\{(1,2),(3)\}$
14.	$\{(1),(2)\}$	$\{(1,2),(2,3)\}$
15.	$\{(1),(2)\}$	$\{(1,3),(2,3)\}$
16.	$\{(1),(2)\}$	$\{(1,2),(1,3),(2,3)\}$
17.	$\{(1),(2)\}$	$\{(1,2,3)\}$
18.	$\{(1),(2),(3)\}$	$\{(1,2)\}$
19.	$\{(1),(2),(3)\}$	$\{(1,2),(2,3)\}$
20.	$\{(1),(2),(3)\}$	$\{(1,2),(1,3),(2,3)\}$
21.	$\{(1),(2),(3)\}$	$\{(1,2,3)\}$
22.	$\{(1,2)\}$	$\{(3)\}$
23.	$\{(1,2)\}$	$\{(2,3)\}$
24.	$\{(1,2)\}$	$\{(1,3),(2,3)\}$
25.	$\{(1,2)\}$	$\{(1,2,3)\}$
26.	$\{(1),(2,3)\}$	$\{(1,2)\}$
27.*	$\{(1),(2,3)\}$	$\{(1,2),(1,3)\}$
28.	$\{(1),(2,3)\}$	$\{(1,2,3)\}$
29.	$\{(1,2),(2,3)\}$	$\{(1,3)\}$
30.	$\{(1,2),(2,3)\}$	$\{(1,2,3)\}$
31.	$\{(1,2),(1,3),(2,3)\}$	$\{(1,2,3)\}$



MDCS of Example 2.1.1.



MDCS of Example 2.1.2.

Figure 2.2: MDCS of Example 2.1.1 and Example 2.1.2

It can be shown that superposition is optimal in most of the cases above. The proof can be done by constructing the $\mathbf{R_{sp}}$ and proving the converse of the coding theorem. Only two cases fail in the coverse and they are picked out and another optimal coding scheme is identified and the admissible coding rate regions are found.

The coding rate region induced by superposition $\mathbf{r_{sp}}$ is evident once the fan of the decoders are known. The method to construct the equivalent $\mathbf{R_{sp}}$ from $\mathbf{r_{sp}}$ is elaborated in Chapter 4. Here we just give $\mathbf{r_{sp}}$ and $\mathbf{R_{sp}}$ for 6 of the 25 cases. Now for the two cases for which superposition is not optimal, the admissible coding rate region are found by first locating all the extreme coding rate tuples and then finding out the constraints that intersect at those points. (It is a method of 'brute force' , but luckily we are able to solve the problem this way.) We will give both the coding rate region induced by superposition and the general admissible coding rate region for comparison.

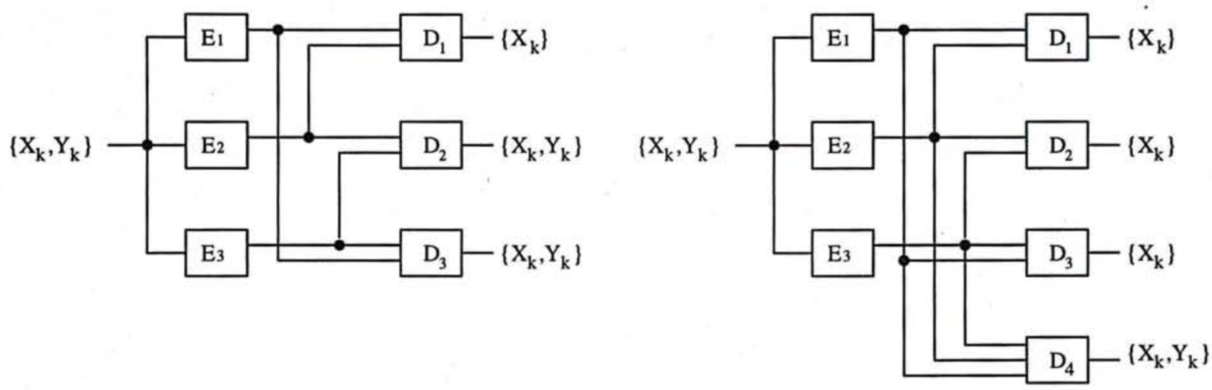
Example 2.1.1 case (15)

The MDCS is shown in Figure 2.2. The coding rate region $\mathbf{r_{sp}}$ is given by $\{(R_1, R_2, R_3): \ R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_1^x &\geq H(X) \\ r_2^x &\geq H(X) \\ r_1^y + r_3^y &\geq H(Y) \\ r_2^y + r_3^y &\geq H(Y) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by $\{(R_1, R_2, R_3): \ R_i \geq 0$ for $i = 1, 2, 3$ and

$$R_1 \geq H(X)$$



MDCS of Example 2.1.3. MDCS of Example 2.1.4.

Figure 2.3: MDCS of Example 2.1.3 and 2.1.4

$$
\begin{aligned}
R_2 &\geq H(X) \\
R_1 + R_3 &\geq H(X) + H(Y) \\
R_2 + R_3 &\geq H(X) + H(Y) \qquad \}
\end{aligned}
$$

Example 2.1.2 case (16)
The MDCS is shown in Figure 2.2. The coding rate region $\mathbf{r_{sp}}$ is given by $\{(R_1, R_2, R_3): \quad R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
r_1^x &\geq H(X) \\
r_2^x &\geq H(X) \\
r_1^y + r_2^y &\geq H(Y) \\
r_1^y + r_3^y &\geq H(Y) \\
r_2^y + r_3^y &\geq H(Y) \qquad \}
\end{aligned}
$$

The equivalent $\mathbf{R_{sp}}$ is given by $\{(R_1, R_2, R_3): \quad R_i \geq 0$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
R_1 &\geq H(X) \\
R_2 &\geq H(X) \\
R_1 + R_3 &\geq 2H(X) + H(Y) \\
R_1 + R_3 &\geq H(X) + H(Y) \\
R_2 + R_3 &\geq H(X) + H(Y) \qquad \}
\end{aligned}
$$

Example 2.1.3 case (24)
The MDCS is shown in Figure 2.3. The coding rate region $\mathbf{r_{sp}}$ is given by $\{(R_1, R_2, R_3): \quad R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i = 1, 2, 3$ and

$$
r_1^x + r_2^x \geq H(X)
$$

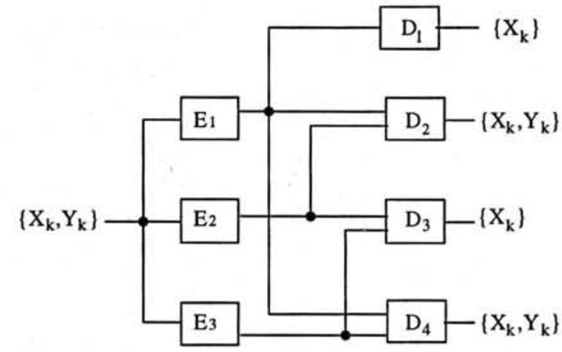


Figure 2.4: MDCS of Example 2.1.5

$$\begin{aligned} r_1^y + r_3^y &\geq H(Y) \\ r_2^y + r_3^y &\geq H(Y) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by $\{(R_1, R_2, R_3): \ R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 + R_2 &\geq H(X) \\ R_1 + R_3 &\geq H(X) + H(Y) \\ R_2 + R_3 &\geq H(X) + H(Y) \qquad \} \end{aligned}$$

Example 2.1.4 case (31)
The MDCS is shown in Figure 2.3. The coding rate region $\mathbf{r_{sp}}$ is given by $\{(R_1, R_2, R_3): \ R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_1^x + r_2^x &\geq H(X) \\ r_1^x + r_3^x &\geq H(X) \\ r_2^x + r_3^x &\geq H(X) \\ r_1^y + r_2^y + r_3^y &\geq H(Y) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by $\{(R_1, R_2, R_3): \ R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 + R_2 &\geq H(X) \\ R_1 + R_3 &\geq H(X) \\ R_2 + R_3 &\geq H(X) \\ R_1 + R_2 + R_3 &\geq \frac{3}{2}H(X) + H(Y) \qquad \} \end{aligned}$$

Example 2.1.5 case (27)
This is a case where an optimal rate is not achieved by superposition but by linear combination of streams.

The configuration of the MDCS is shown in Figure 2.4. The $\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, R_3): \quad R_i = r_i^x + r_i^y$ where $r_i^x, r_i^y \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_1^x &\geq H(X) \\ r_2^x + r_3^x &\geq H(X) \\ r_1^y + r_2^y &\geq H(Y) \\ r_1^y + r_3^y &\geq H(Y) \quad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by
$\{(R_1, R_2, R_3): \quad R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 + R_3 &\geq H(X) \\ R_1 + R_3 &\geq H(X) + H(Y) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ 2R_1 + R_2 + R_3 &\geq 3H(X) + 2H(Y) \quad \} \end{aligned}$$

Firstly let us see why superposition is not optimal. If we assume that $\{X_k\}, \{Y_k\}$ are independent bit streams with

$$p[X_k = 1] = p[X_k = 0] = p[Y_k = 1] = p[Y_k = 0] = 0.5$$

So $H(X) = H(Y) = 1$, the rate tuple $(1, 1, 1)$ is admissible by coding $\{(X_k, Y_k)\}$ in encoder 1, 2 and 3 as $\{X_k\}, \{Y_k\}, \{X_k \oplus Y_k\}$ respectively. But this tuple is not in $\mathbf{R_{sp}}$ since it does not satisfy the last contraint inequality in $\mathbf{R_{sp}}$.

Now we state the admissible coding rate region. First it is a standard argument that if all the extreme rate tuples of a coding rate region is admissible then the whole coding rate region is admissible. So the forward part of the coding theorem is proved by showing all the extereme tuples of the region are admissible. Then converse part is proved in steps similiar to Example 1.1 in Chapter 1. It is obvious so the steps are not shown. The admissible coding rate region is
$\{(R_1, R_2, R_3): \quad R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 + R_3 &\geq H(X) \\ R_1 + R_3 &\geq H(X) + H(Y) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) \quad \} \end{aligned}$$

We are to examine the extreme tuples of the above region. The problem turns out to be very subtle here. The set of extreme tuples in the case $H(X) > H(Y)$ are different from those in the case $H(X) < H(Y)$. We examine both cases separately. The admissible coding rate regions for both cases are shown in Figure 2.5.

a) $H(X) > H(Y)$

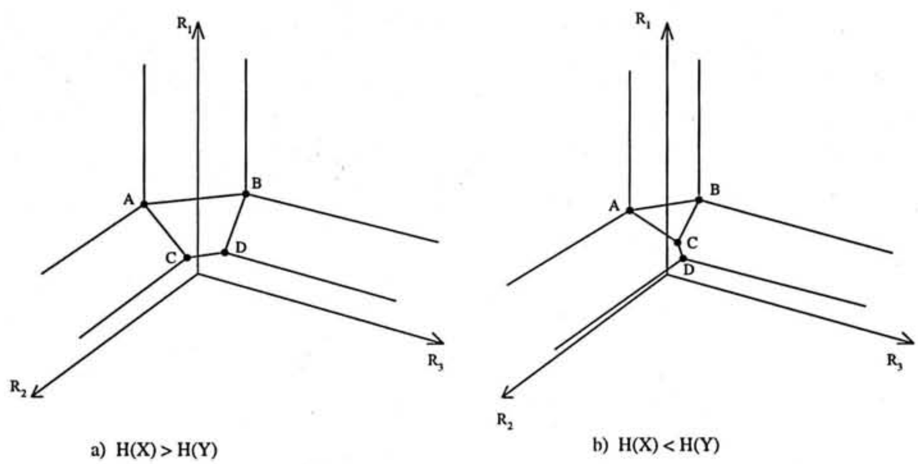


Figure 2.5: Admissible coding rate region for the MDCS in Example 2.1.5

There are four extreme tuples in this region. They are found to be

$$\begin{aligned} A &= (H(X)+H(Y), H(X), 0) \\ B &= (H(X)+H(Y), 0, H(X)) \\ C &= (H(X), H(X), H(Y)) \\ D &= (H(X), H(Y), H(X)) \end{aligned}$$

The first two tuples are clearly admissible. We define $\{X_k\} = \{(X_k^1, X_k^2)\}$ where $\{Y_k\}, \{X_k^1\}$ and $\{X_k^2\}$ are independent and $\{X_k^1\}$ has the same entropy rate as $\{Y_k\}$. (In physical bit stream operation, it means dividing the output bits per unit time of the streams represented by X into two independent parts represented by X^1 and X^2, with the first part having $H(Y)$ bits). It is feasible to code the 2 streams as

$$\{X_k^1, X_k^2\}, \{Y_k, X_k^2\} and \{X_k^1 \oplus Y_k\}$$

or

$$\{X_k^1, X_k^2\}, \{Y_k\} and \{X_k^1 \oplus Y_k, X_k^2\}$$

in encoder 1, 2 and 3 respectively. Now,

$$\begin{aligned} H(Y, X^2) &= H(Y) + H(X^2) = H(X^1) + H(X^2) = H(X) \\ H(X^1 \oplus Y) &= max\{H(X^1), H(Y)\} = H(X^1) = H(Y) \\ H(X^1 \oplus Y, X^2) &= H(X^1 \oplus Y) + H(X^2) = H(X^1) + H(X^2) = H(X) \end{aligned}$$

So the last two extreme triples corresponding to the two coding schemes above are admissible.

b) $H(X) < H(Y)$

There are also four extreme tuples in this region. They are found to be

$$\begin{aligned} A &= (H(X)+H(Y), H(Y), 0) \\ B &= (H(X)+H(Y), 0, H(Y)) \end{aligned}$$

$$\begin{aligned} C &= (H(Y), H(X), H(X)) \\ D &= (H(X), H(Y), H(Y)) \end{aligned}$$

The first two tuples are clearly admissible. We define $\{Y_k\} = \{(Y_k^1, Y_k^2)\}$ where $\{X_k\}, \{Y_k^1\}$ and $\{Y_k^2\}$ are independent and $\{Y_k^1\}$ has the same entropy rate as $\{X_k\}$ and it is feasible to code the 2 streams as

$$\{X_k, Y_k^2\}, \{Y_k^1\} and \{X_k \oplus Y_k^1\}$$

or

$$\{X_k\}, \{Y_k^1, Y_k^2\} and \{X_k \oplus Y_k^1, Y_k^2\}$$

in encoder 1, 2 and 3 respectively. Now,

$$\begin{aligned} H(Y_1, Y_2) &= H(Y) \\ H(X, Y^2) &= H(X) + H(Y^2) = H(Y^1) + H(Y^2) = H(Y) \\ H(X \oplus Y^1) &= max\{H(X), H(Y^1)\} = H(Y^1) = H(X) \\ H(X \oplus Y^1, Y^2) &= H(X \oplus Y^1) + H(Y^2) = H(Y^1) + H(Y^2) = H(Y) \end{aligned}$$

So the last two extreme triples corresponding to the two coding schemes above are admissible.

If $H(X) = H(Y)$, the last two extreme rate triples is identical and it is simply achieved by coding the two streams as $\{X_k\}, \{Y_k\}$ and $\{X_k \oplus Y_k\}$ in encoder 1, 2 and 3 repectively. Now there are only three extreme tuples.

Note that the MDCS in this example not a sequentially refinable MDCS as explained in the last section of Chapter 1 but it belongs to a more general class of MDCS which apply linear combination and will be discussed in Appendix B.

Example 2.1.6 case (8)

This example is the same as Example 1.2 given in Chapter 1 and $\mathbf{R_{sp}}$ has already been stated. We just state the admissible coding rate region here. Again we assume $H(X) = H(Y)$.

$\{(R_1, R_2, R_3):\ R_i \geq 0$ for $i = 1, 2, 3$ and

$$R_1 \geq H(X) \tag{2.1}$$
$$R_2 + R_3 \geq H(X) + H(Y) \tag{2.2}$$
$$R_1 + R_3 \geq H(X) + H(Y) \tag{2.3}$$
$$R_1 + R_2 \geq H(X) + H(Y) \quad \} \tag{2.4}$$

Just like Example 2.1.5, the set of extreme points of the above region in the case $H(X) > H(Y)$ are different from those in the case $H(X) < H(Y)$. We examine both cases separately.

a) $H(X) > H(Y)$

There are four extreme tuples in this region. They are found to be

$$\begin{aligned} &(H(X) + H(Y), H(X) + H(Y), 0) \\ &(H(X) + H(Y), 0, H(X) + H(Y)) \end{aligned}$$

$$(H(X), H(X), H(Y))$$
$$(H(X), H(Y), H(X))$$

The first two tuples are clearly admissible. We define $\{X_k\} = \{(X_k^1, X_k^2)\}$ where $\{Y_k\}, \{X_k^1\}$ and $\{X_k^2\}$ are independent and $\{X_k^1\}$ has the same entropy rate as $\{Y_k\}$. It is feasible to code the 2 streams as $\{X_k^1, X_k^2\}, \{Y_k, X_k^2\} and \{X_k^1 \oplus Y_k, X_k^2\}$ or $\{X_k^1, X_k^2\}, \{Y_k, X_k^2\} and \{X_k^1 \oplus Y_k\}$ in encoder 1, 2 and 3 repectively. By the arguments in the Example 2.1.5, the last two extreme triples corresponding to the two coding schemes above are admissible.

b) $H(X) < H(Y)$

There are also four extreme tuples in this region. They are found to be

$$\begin{aligned}
&(H(X) + H(Y), H(X) + H(Y), 0) \\
&(H(X) + H(Y), 0, H(X) + H(Y)) \\
&(\frac{H(Y) + H(X)}{2}, \frac{H(Y) + H(X)}{2}, \frac{H(Y) + H(X)}{2}) \\
&(H(X), H(Y), H(Y))
\end{aligned}$$

The first two tuples are clearly admissible. We define $\{Y_k\} = \{(Y_k^1, Y_k^2)\}$ where $\{X_k\}, \{Y_k^1\}$ and $\{Y_k^2\}$ are independent and $\{Y_k^1\}$ has the same entropy rate as $\{X_k\}$, then we define $\{Y_k^2\} = \{(Y_k^{21}, Y_k^{22})\}$ where $\{X_k\}, \{Y_k^2\}, \{Y_k^{21}\}$ and $\{Y_k^{22}\}$ are independent and $H(Y^{21}) = H(Y^{22})$. It is feasible to code the two streams in encoder 1, 2 and 3 as

$$\{X_k, Y_k^{21}\}, \{Y_k^1, Y_k^{22}\} and \{X_k \oplus Y_k^1, Y_k^{21} \oplus Y_k^{22}\}$$

or

$$\{X_k\}, \{Y_k^1, Y_k^2\} and \{X_k \oplus Y_k^1, Y_k^2\}$$

Now,

$$\begin{aligned}
H(Y_1, Y_2) &= H(Y) \\
H(Y^{21}) &= H(Y^{22}) = \frac{H(Y^2)}{2} = \frac{H(Y) - H(Y^1)}{2} = \frac{H(Y) - H(X)}{2} \\
H(X, Y^{21}) &= H(X) + H(Y^{21}) = H(X) + \frac{H(Y) - H(X)}{2} = \frac{H(Y) + H(X)}{2} \\
H(Y_1, Y^{21}) &= H(Y_1) + H(Y^{21}) = H(X) + H(Y^{21}) = \frac{H(Y) + H(X)}{2} \\
H(X \oplus Y^1) &= max\{H(X), H(Y^1)\} = H(Y^1) = H(X) \\
H(Y^{21} \oplus Y^{22}) &= max\{H(Y^{21}), H(Y^{22})\} = H(Y^{21}) = \frac{H(Y) - H(X)}{2} \\
H(X \oplus Y^1, Y^{21} \oplus Y^{22}) &= H(X \oplus Y^1) + H(Y^{21} \oplus Y^{22}) \\
&= H(X) + \frac{H(Y) - H(X)}{2} = \frac{H(Y) + H(X)}{2} \\
H(X \oplus Y^1, Y^2) &= H(X \oplus Y^1) + H(Y^2) = H(Y^1) + H(Y^2) = H(Y)
\end{aligned}$$

So the last two extreme triples corresponding to the two coding schemes above are admissible.

For the case $H(X) = H(Y)$, the last two extreme tuples become identical and so there are only three extreme tuples for the admissible coding rate region.

It is also very obvious that any admissible tuples (R_1, R_2, R_3) satisfies all the rate constraints of the coding rate region.

This example is a member of the class of *sequentially refinable* MDCS introduced by Roche [12]. Also it belongs to the class of MDCS's discussed in Appendix B for which superposition is always not optimal.

The last two examples is very important as we will see that they are the common *embedded* MDCS's (expained in next section) of all cases of 3-level-3-encoder MDCS's for which superposition is not optimal.

Here we conclude this section by stating that superposition is optimal in all but two cases in 2-level-3-encoder MDCS's. Superposition and linear combination of sources are the *only two* complementary optimal coding schemes.

2.2 3-level-3-encoder MDCS

Now we continue to enumerate the possible structures for a 3-level MDCS with 3 encoders. No level-3 decoder can be added onto a 2-level MDCS if the fan of a level-2 decoder has already covered all the encoders in the MDCS's since the fan of a level-3 decoder can neither be identical to the fan of any level-2 decoder nor a subset of it. So those MDCS's in Section 2.1 with fan of a level-2 decoder as $\{(1, 2, 3)\}$ are excluded in further consideration (totally 7 of them). Then we add all the possible fans of level-3 decoders onto the rest 24 cases to construct 3-level MDCS's. Totally 69 different structures appear and are listed in the following. (*) indicates that superposition is not optimal in the MDCS. Again symmetrical structures are treated as identical.

	level 1	*level* 2	*level* 3
1.	$\{(1)\}$	$\{(2)\}$	$\{3\}$
2.	$\{(1)\}$	$\{(2)\}$	$\{(1,2)\}$
3.	$\{(1)\}$	$\{(2)\}$	$\{(1,3)\}$
4.	$\{(1)\}$	$\{(2)\}$	$\{(2,3)\}$
5.	$\{(1)\}$	$\{(2)\}$	$\{(1,2),(1,3)\}$
6.	$\{(1)\}$	$\{(2)\}$	$\{(1,2),(2,3)\}$
7.	$\{(1)\}$	$\{(2)\}$	$\{(1,3),(2,3)\}$
8.*	$\{(1)\}$	$\{(2)\}$	$\{(1,2),(1,3),(2,3)\}$
9.	$\{(1)\}$	$\{(2)\}$	$\{(1,2,3)\}$
10.	$\{(1)\}$	$\{(2),(3)\}$	$\{(1,2)\}$
11.	$\{(1)\}$	$\{(2),(3)\}$	$\{(2,3)\}$
12.	$\{(1)\}$	$\{(2),(3)\}$	$\{(1,3),(2,3)\}$
13.	$\{(1)\}$	$\{(2),(3)\}$	$\{(1,2),(1,3)\}$
14.	$\{(1)\}$	$\{(2),(3)\}$	$\{(1,2),(1,3),(2,3)\}$
15.	$\{(1)\}$	$\{(2),(3)\}$	$\{(1,2,3)\}$
16.	$\{(1)\}$	$\{(1,2)\}$	$\{(3)\}$
17.	$\{(1)\}$	$\{(1,2)\}$	$\{(2,3)\}$
18.	$\{(1)\}$	$\{(1,2)\}$	$\{(1,3)\}$
19.	$\{(1)\}$	$\{(1,2)\}$	$\{(1,2,3)\}$
20.*	$\{(1)\}$	$\{(1,2)\}$	$\{(1,3),(2,3)\}$
21.	$\{(1)\}$	$\{(1,2),(3)\}$	$\{(1,3)\}$
22.	$\{(1)\}$	$\{(1,2),(3)\}$	$\{(2,3)\}$
23.*	$\{(1)\}$	$\{(1,2),(3)\}$	$\{(1,3),(2,3)\}$
24.	$\{(1)\}$	$\{(1,2),(3)\}$	$\{(1,2,3)\}$
25.	$\{(1)\}$	$\{(2,3)\}$	$\{(1,2)\}$
26.*	$\{(1)\}$	$\{(2,3)\}$	$\{(1,2),(1,3)\}$
27.	$\{(1)\}$	$\{(2,3)\}$	$\{(1,2,3)\}$
28.*	$\{(1)\}$	$\{(1,2),(2,3)\}$	$\{(1,3)\}$

29.	$\{(1)\}$	$\{(1,2),(2,3)\}$	$\{(1,2,3)\}$
30.*	$\{(1)\}$	$\{(1,2),(1,3)\}$	$\{(2,3)\}$
31.	$\{(1)\}$	$\{(1,2),(1,3)\}$	$\{(1,2,3)\}$
32.*	$\{(1)\}$	$\{(1,2),(1,3),(2,3)\}$	$\{(1,2,3)\}$
33.	$\{(1),(2)\}$	$\{(3)\}$	$\{(1,2)\}$
34.	$\{(1),(2)\}$	$\{(3)\}$	$\{(2,3)\}$
35.	$\{(1),(2)\}$	$\{(3)\}$	$\{(1,2),(2,3)\}$
36.	$\{(1),(2)\}$	$\{(3)\}$	$\{(1,3),(2,3)\}$
37.*	$\{(1),(2)\}$	$\{(3)\}$	$\{(1,2),(1,3),(2,3)\}$
38.	$\{(1),(2)\}$	$\{(3)\}$	$\{(1,2,3)\}$
39.	$\{(1),(2)\}$	$\{(1,2)\}$	$\{(3)\}$
40.	$\{(1),(2)\}$	$\{(1,2)\}$	$\{(2,3)\}$
41.	$\{(1),(2)\}$	$\{(1,2)\}$	$\{(1,2,3)\}$
42.	$\{(1),(2)\}$	$\{(1,2)\}$	$\{(1,3),(2,3)\}$
43.	$\{(1),(2)\}$	$\{(1,2),(3)\}$	$\{(2,3)\}$
44.*	$\{(1),(2)\}$	$\{(1,2),(3)\}$	$\{(1,3),(2,3)\}$
45.	$\{(1),(2)\}$	$\{(1,2),(3)\}$	$\{(1,2,3)\}$
46.	$\{(1),(2)\}$	$\{(2,3)\}$	$\{(1,2)\}$
47.	$\{(1),(2)\}$	$\{(2,3)\}$	$\{(1,2),(1,3)\}$
48.	$\{(1),(2)\}$	$\{(2,3)\}$	$\{(1,2,3)\}$
49.	$\{(1),(2)\}$	$\{(1,2),(1,3)\}$	$\{(2,3)\}$
50.	$\{(1),(2)\}$	$\{(1,2),(1,3)\}$	$\{(1,2,3)\}$
51.	$\{(1),(2)\}$	$\{(1,3),(2,3)\}$	$\{(1,2)\}$
52.	$\{(1),(2)\}$	$\{(1,3),(2,3)\}$	$\{(1,2,3)\}$
53.	$\{(1),(2)\}$	$\{(1,2),(1,3),(2,3)\}$	$\{(1,2,3)\}$
54.	$\{(1),(2),(3)\}$	$\{(1,2)\}$	$\{(2,3)\}$
55.	$\{(1),(2),(3)\}$	$\{(1,2)\}$	$\{(1,3),(2,3)\}$
56.	$\{(1),(2),(3)\}$	$\{(1,2)\}$	$\{(1,2,3)\}$
57.	$\{(1),(2),(3)\}$	$\{(1,2),(2,3)\}$	$\{(1,3)\}$
58.	$\{(1),(2),(3)\}$	$\{(1,2),(2,3)\}$	$\{(1,2,3)\}$
59.	$\{(1),(2),(3)\}$	$\{(1,2),(1,3),(2,3)\}$	$\{(1,2,3)\}$
60.	$\{(1,2)\}$	$\{(3)\}$	$\{(2,3)\}$

61.*	$\{(1,2)\}$	$\{(3)\}$	$\{(1,3),(2,3)\}$
62.	$\{(1,2)\}$	$\{(3)\}$	$\{(1,2,3)\}$
63.	$\{(1,2)\}$	$\{(2,3)\}$	$\{(1,3)\}$
64.	$\{(1,2)\}$	$\{(2,3)\}$	$\{(1,2,3)\}$
65.	$\{(1,2)\}$	$\{(1,3),(2,3)\}$	$\{(1,2,3)\}$
66.*	$\{(1),(2,3)\}$	$\{(1,2)\}$	$\{(1,3)\}$
67.	$\{(1),(2,3)\}$	$\{(1,2)\}$	$\{(1,2,3)\}$
68.*	$\{(1),(2,3)\}$	$\{(1,2),(1,3)\}$	$\{(1,2,3)\}$
69.	$\{(1,2),(2,3)\}$	$\{(1,3)\}$	$\{(1,2,3)\}$

For this class of MDCS, *twelve* cases for which superposition is not optimal are discovered. All the twelve cases has one crucial similarity in structure and are explained below. Here we define a new term to facilitate further discussion.

Definition

An *embedded MDCS* of an MDCS A is the MDSC obtained by degenerating some of the data streams encoded in MDCS A.

Example 2.2.1

Consider the 3-level-3-encoder-MDCS in case (61) above, where streams $\{X_k\}$, $\{(X_k, Y_k)\}$, and $\{(X_k, Y_k, Z_k)\}$ are recovered from level-1, level-2 and level-3 decoders respectively. The fan of the decoders of the three different levels are as follows.

level 1	*level* 2	*level* 3
$\{(1,2)\}$	$\{(3)\}$	$\{(1,3),(2,3)\}$

Three embedded MDCS's (2-level) are obtained by degenerating streams $\{X_k\}$, $\{Y_k\}$ and $\{Z_k\}$ respectively and are listed below by showing the structure of the fan of their level-1 and level-2 decoders.

	level 1	*level* 2
1.	$\{(3)\}$	$\{(1,3),(2,3)\}$
2.	$\{(1,2),(3)\}$	$\{(1,3),(2,3)\}$
3.	$\{(1,2)\}$	$\{(3)\}$

Note that embedded MDCS's exist only for MDCS's with more than 2 levels of decoders. Also the original MDCS can be constructed from the embedded MDCS by adding one level of decoders which can recover the data stream which is degenerated when the MDCS is reduced to the embedded MDCS. We will show that the admissible coding rate region of the MDCS can also be constructed with the knowledge of the admissible coding rate region of the embedded MDCS.

We discovered that *all 3-level-3-encoder MDCS's for which superposition is not optimal has an embedded MDCS which is one of the two 2-level-3-encoder MDCS's for which superposition is not optimal.*

If an admissible rate tuple of the embedded MDCS of an MDCS cannot be achieved by superposition, then an admissible rate tuple of the original MDCS is still not achievable by superposition when the rate of the missing stream is added. In Example 2.2.1 (case (61)), the streams $\{(X_k, Z_k)\}$ in the embedded MDCS (2) above can be coded in encoder 1, 2 and 3 as $\{X_k \oplus Z_k\}$, $\{Z_k\}$ and $\{X_k\}$ respectively and the rate is $(H(X), H(Z), H(X))$ (assume $H(X) = H(Z)$ for simplicity). It is shown in Example 2.1.5 (an MDCS symmetric to embedded MDCS (2) above) that this tuple is not achieved by superposition. And one of the admissible rate tuple for the original MDCS in the example is $(H(X), H(Z), H(X) + H(Y))$ which is obtained by adding the rate of encoding $\{Y_k\}$ onto the tuple $(H(X), H(Z), H(X))$ such that it is sufficient for the recovery of $\{Y_k\}$ by various level-3 and level-2 decoders of the original MDCS. This rate tuple also is not obtained by superimposing the 3 streams together. So superposition is not optimal in the MDCS in Example 2.2.1 because it is not optimal in one of its embedded MDCS.

We can also extend the discussion to an m-level MDCS (denoted it as A). By the principle of superposition, coding rate of an encoder is the aggregation of coding rate of all the independent data streams in MDCS A and it is always feasible to do so. Suppose in MDCS A, independent data streams $\{(X_k^1, X_k^2, \ldots, X_k^m)\}$ are encoded. If an embedded MDCS in A is obtained by degerating all the streams except $\{X_k^{i_1}, X_k^{i_2}, \ldots, X_k^{i_j}\}$ where $1 \leq i_1 < i_2 < \ldots < i_j \leq m$ and they can be coded by another optimal scheme in this embedded MDCS and the rate is not achievable by superposition. Then an admissible coding rate tuple for MDCS A is that the rate for those streams $\{X_k^{i_1}, \ldots, X_k^{i_j}\}$ can be replaced by the optimal rates of the alternative coding scheme while the admissible rates for other streams remains unchanged.

Obviously that coding rate cannot be achieved by superposition. So superposition is not optimal in MDCS A since superposition is not optimal in one of the embedded MDCS's. So as a conclusion one of the methods to examine the optimality of superposition in a general m-level MDCS is as follows:

In an MDCS, if any embedded MDCS of 2 or more levels can be replaced by an alternative coding scheme which can achieve a lower coding rate than the scheme applying superposition, then the superposition is not an optimal coding scheme for the MDCS.

Also, we have a high degree of belief that superposition is optimal if it is optimal in all embedded MDCS's of the MDCS since the only possibility left is to code all the streams altogether in a single level to achieve a rate lower than that achieved by superposition.

Embedded MDCS's are usually relatively simpler than the original MDCS. It is quite difficult to characterize the admissible coding rate region for the original MDCS, especially when superposition is not optimal. But characterizing the region of the embedded MDCS is relatively easier (though it is still nontrivial). In the 3-level-3-encoder MDCS's, all the possible embedded MDCS's must be 2-level-3-encoder MDCS's and for those cases for which superposition is not optimal, the common embedded MDCS is either case (8) or case (27) in the 2-level-encoder-MDCS. Luckily we can characterize the whole admissible coding rate regions for two cases above. We can superimpose the admissible coding rate region for the missing stream onto the that of the two 2-level cases and construct a feasible coding rate region for the twelve 3-level-3-encoder MDCS's for which superposition is not optimal. The resulting region must be a subset of the admissible coding rate region. The method is similar to that of constructing $\mathbf{R}_{\mathbf{sp}}$ from $\mathbf{r}_{\mathbf{sp}}$.

Surprisingly, it turns out that the coding rate region formed by adding the rate of the missing data stream onto the admissible region of the embedded MDCS is the general admissible coding rate region. (See Example 2.2.1 to Example 2.2.4) This discovery implies that the optimal coding scheme for the twelve 3-level cases is found to be *neither* superposition *nor* linear combination of streams, but a *hybrid* of the two.

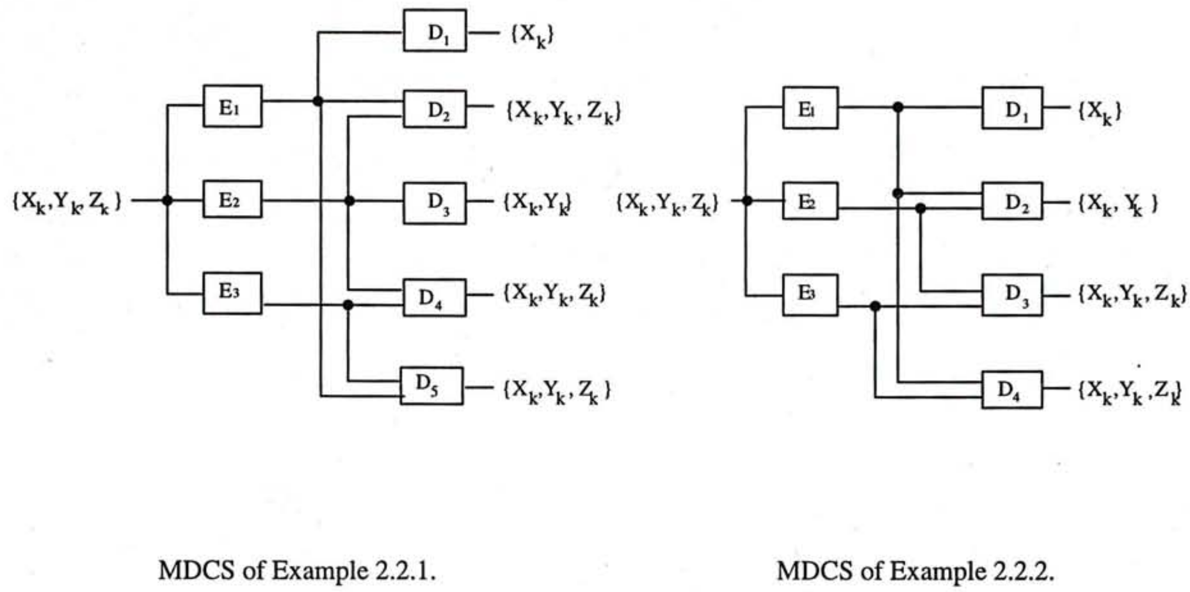


Figure 2.6: MDCS of Example 2.2.1 and Example 2.2.2

Here we can conclude that for 3-level-3-encoder MDCSs', optimal coding scheme is either superposition or a hybrid of superposition and linear combination of sources.

In the following we list the admissible coding rate region for four of the twelve cases for which superposition is not optimal. The extreme tuples in the regions are listed so that the forward part of the coding theorem becomes obvious. The coverse part can be quite easily proved by interested readers. We also indicate how to code the sources in the embedded MDCS's. We also give the coding rate region in terms of the subrate constraints. In all the cases, we assume that the two data streams coded in the embedded MDCS's have the *same* entropy rate for simplicity. For the general cases, the problems can be analyzed in the same spirit as in Example 2.1.5 and Example 2.1.6 and are omitted here.

Example 2.2.1 (case 8)

The coding rate region induced by superimposing the rate of stream $\{X_k\}$ onto that of the embedded MDCS in which $\{Y_k\}$ and $\{Z_k\}$ are coded is given in the following, where r_i^{yz} represents the rate of the embedded MDCS and r_i^x repersents the rate of stream $\{X_k\}$ in encoder i.

$\{(R_1, R_2, R_3):\ \ R_i = r_i^{yz} + r_i^x$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_i^{yz},\ r_i^x &\geq 0 \text{ for } i = 1,2,3 \\ r_1^x &\geq H(X) \\ r_2^x &\geq H(X) \\ r_2^{yz} &\geq H(Y) \\ r_1^{yz} + r_2^{yz} &\geq H(Y) + H(Z) \end{aligned}$$

$$
\begin{aligned}
r_2^{yz} + r_3^{yz} &\geq H(Y) + H(Z) \\
r_1^{yz} + r_3^{yz} &\geq H(Y) + H(Z) \qquad \}
\end{aligned}
$$

One may notice that the constraints in terms of r_i^{yz} has the same structurue as that of Example 2.1.6. since the embedded MDCS is of the same structrue as the MDCS of Example 2.1.6.

The admissible coding rate region, which is equivalent to the above one, is given by $\{(R_1, R_2, R_3): \; R_i \geq 0$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
R_1 &\geq H(X) \\
R_2 &\geq H(X) + H(Y) \\
R_1 + R_2 &\geq 2H(X) + H(Y) + H(Z) \\
R_1 + R_3 &\geq H(X) + H(Y) + H(Z) \\
R_2 + R_3 &\geq H(X) + H(Y) + H(Z) \qquad \}
\end{aligned}
$$

The extreme tuples are

$$
\begin{gathered}
(H(X) + H(Y) + H(Z), H(X) + H(Y) + H(Z), 0) \\
(H(X), H(X) + H(Y) + H(Z), H(Y) + H(Z)) \\
(H(X) + H(Z), H(X) + H(Y), H(Z))
\end{gathered}
$$

The last extreme tuple is admissible by coding the stream $\{Y_k\}$ and $\{Z_k\}$ in encoder 1, 2 and 3 as $\{Z_k\}$, $\{Y_k\}$ and $\{Y_k \oplus Z_k\}$ respectively.

Example 2.2.2 (case 19)

The coding rate region induced by superimposing the rate of stream $\{Z_k\}$ onto that of the embedded MDCS in which $\{X_k\}$ and $\{Y_k\}$ are coded is given in the following, where r_i^{xy} represents the rate of the embedded MDCS, r_i^z repersents the rate of stream $\{Z_k\}$ in encoder i. The admissible coding rate region is

$\{(R_1, R_2, R_3): \; R_i = r_i^{xy} + r_i^z \geq 0$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
r_i^{xy},\; r_i^z &\geq 0 \text{ for } i = 1, 2, 3 \\
r_1^{xy} &\geq H(X) \\
r_1^{xy} + r_2^{xy} &\geq H(X) + H(Y) \\
r_2^{xy} + r_3^{xy} &\geq H(X) + H(Y) \\
r_1^{yz} + r_3^{xy} &\geq H(X) + H(Y) \\
r_1^z + r_3^z &\geq H(Z) \\
r_2^z + r_3^z &\geq H(Z) \qquad \}
\end{aligned}
$$

One may notice that the constraints in terms of r_i^{xy}has the same structurue as that of Example 2.1.6. since the embedded MDCS is of the same structrue as the MDCS of Example 2.1.6. The admissible coding rate region, which is equivalent to the above one, is given by

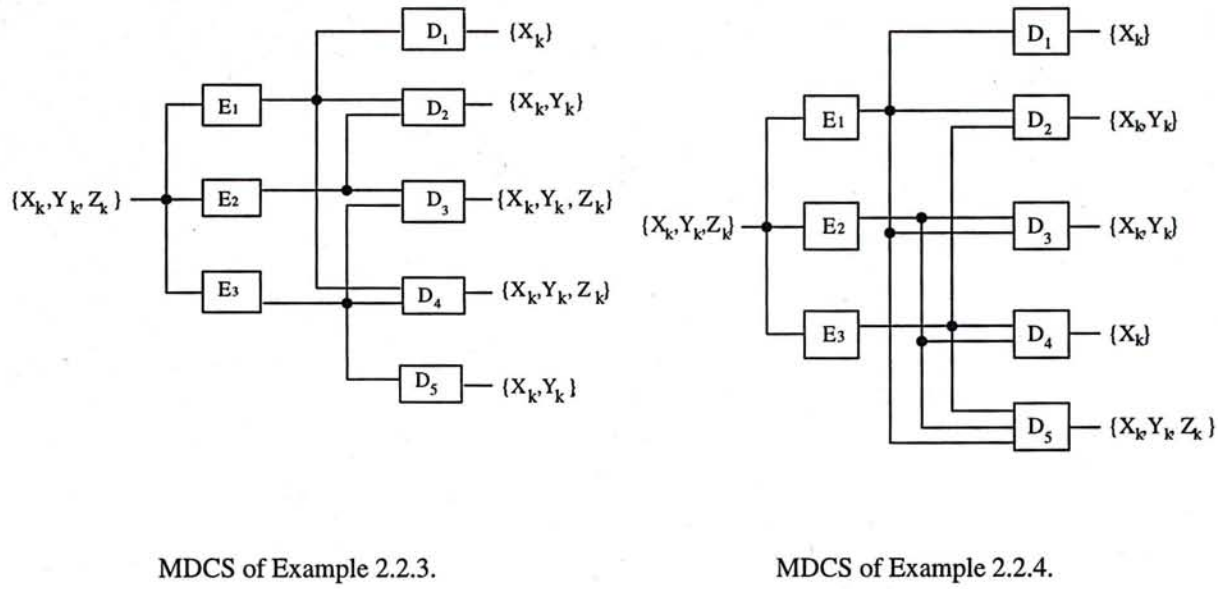


Figure 2.7: MDCS of Example 2.2.3 and Example 2.2.4

$\{(R_1, R_2, R_3): \quad R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ R_1 + R_3 &\geq H(X) + H(Y) + H(Z) \\ R_2 + R_3 &\geq H(X) + H(Y) + H(Z) \qquad \} \end{aligned}$$

The extreme tuples are

$$\begin{gathered} (H(X) + H(Y), 0, H(X) + H(Y) + H(Z)) \\ (H(X) + H(Y) + H(Z), H(X) + H(Y) + H(Z)) \\ (H(X), H(Y), H(Y) + H(Z)) \end{gathered}$$

The source $\{X_k\}$ and $\{Y_k\}$ can be coded in encoder 1, 2 and 3 as $\{X_k\}$, $\{Y_k\}$ and $\{X_k \oplus Y_k\}$ respectively.

Example 2.2.3 (case 23)

The coding rate region induced by superimposing the rate of stream $\{X_k\}$ onto that of the embedded MDCS in which $\{Y_k\}$ and $\{Z_k\}$ are coded is given in the following, where r_i^{yz} represents the rate of the embedded MDCS, r_i^x repersents the rate of stream $\{X_k\}$ in encoder i.

$\{(R_1, R_2, R_3): \quad R_i = r_i^{yz} + r_i^x \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_i^{yz},\ r_i^x &\geq 0 \text{ for } i = 1, 2, 3 \\ r_1^x &\geq H(X) \\ r_3^x &\geq H(X) \\ r_3^{yz} &\geq H(Y) \end{aligned}$$

$$
\begin{aligned}
r_1^{yz} + r_2^{yz} &\geq H(Y) \\
r_2^{yz} + r_3^{yz} &\geq H(Y) + H(Z) \\
r_1^{yz} + r_3^{yz} &\geq H(Y) + H(Z) \\
r_1^{yz} + r_2^{yz} + r_3^{yz} &\geq 2H(Y) + H(Z) \qquad \}
\end{aligned}
$$

One may notice that the constraints in terms of r_i^{zy}has the same structurue as that of Example 2.1.5. since the embedded MDCS is of the same structrue as the MDCS of Example 2.1.5.

The admissible coding rate region, which is equivalent to the above one, is given by $\{(R_1, R_2, R_3):\ R_i \geq 0$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
R_1 &\geq H(X) \\
R_3 &\geq H(X) + H(Y) \\
R_1 + R_2 &\geq H(X) + H(Y) \\
R_1 + R_3 &\geq H(X) + H(Y) + H(Z) \\
R_2 + R_3 &\geq H(X) + H(Y) + H(Z) \\
R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \}
\end{aligned}
$$

The extreme tuples are

$$
\begin{gathered}
(H(X) + H(Y), 0, H(X) + H(Y) + H(Z)) \\
(H(X), H(Y), H(Y) + H(Z) + H(X)) \\
(H(Y) + H(X), H(Z), H(X) + H(Y))
\end{gathered}
$$

The source $\{Y_k\}$ and $\{Z_k\}$ can be coded in encoder 1, 2 and 3 as $\{Y_k \oplus Z_k\}$, $\{Z_k\}$ and $\{Y_k\}$ respectively.

Example 2.2.4 (case 68)

This is the same case as Example 1.3. The coding rate region induced by superimposing the rate of stream $\{Z_k\}$ onto that of the embedded MDCS in which $\{X_k\}$ and $\{Y_k\}$ are coded is given in the following, where r_i^{xy} represents the rate of the embedded MDCS, r_i^z repersents the rate of stream $\{Z_k\}$ in encoder i. The admissible coding rate region is

$\{(R_1, R_2, R_3):\ R_i = r_i^{xy} + r_i^z$ for $i = 1, 2, 3$ and

$$
\begin{aligned}
r_i^{xy},\ r_i^z &\geq 0 \text{ for } i = 1, 2, 3 \\
r_1^{xy} &\geq H(X) \\
r_2^{xy} + r_3^{xy} &\geq H(X) \\
r_1^{xy} + r_2^{xy} &\geq H(X) + H(Y) \\
r_1^{xy} + r_3^{xy} &\geq H(X) + H(Y) \\
r_1^{xy} + r_2^{xy} + r_3^{xy} &\geq 2H(X) + H(Y) \\
r_1^z + r_2^z + r_3^z &\geq H(Z) \qquad \}
\end{aligned}
$$

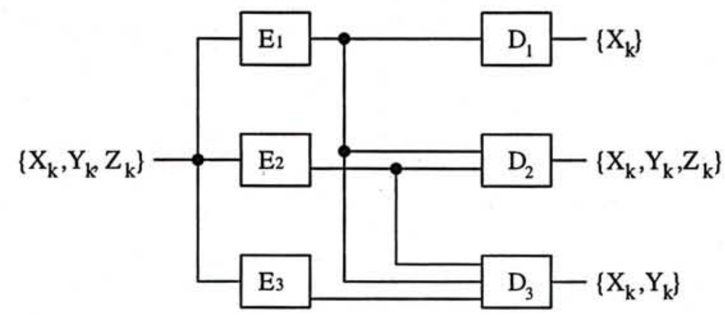


Figure 2.8: MDCS of Example 2.2.5

One may notice that the constraints in terms of r_i^{xy} has the same structurue as that of Example 2.1.5. since the embedded MDCS is of the same structrue as the MDCS of Example 2.1.5. The admissible coding rate region, which is equivalent to the above one, is given by

$\{(R_1, R_2, R_3): \ R_i \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ R_1 + R_3 &\geq H(X) + H(Y) \\ R_2 + R_3 &\geq H(X) \\ R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \} \end{aligned}$$

The extreme tuples are

$$\begin{array}{ll} (H(X)+H(Y)+H(Z), 0, H(X)) & (H(X)+H(Y), 0, H(X)+H(Z)) \\ (H(X)+H(Y)+H(Z), H(X), 0) & (H(X)+H(Y), H(X)+H(Z), 0) \\ (H(X), H(Y), H(X)+H(Z)) & (H(X), H(Y)+H(Z), H(X)) \end{array}$$

The source $\{X_k\}$ and $\{Y_k\}$ can be coded in encoder 1, 2 and 3 as $\{X_k\}$, $\{Y_k\}$ and $\{X_k \oplus Y_k\}$ respectively.

The admissible coding rate regions of four other cases of 3-level-3-encoder MDCS's for which superposition is optimal are also given as examples.

Example 2.2.5 (case 25)

The coding rate region $\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, R_3): \ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3$ and

$$\begin{aligned} r_1^x &\geq H(X) \\ r_2^x + r_3^x &\geq H(X) \\ r_1^y + r_2^y &\geq H(Y) \\ r_2^y + r_3^y &\geq H(Y) \\ r_1^z + r_2^z &\geq H(Z) \qquad \} \end{aligned}$$

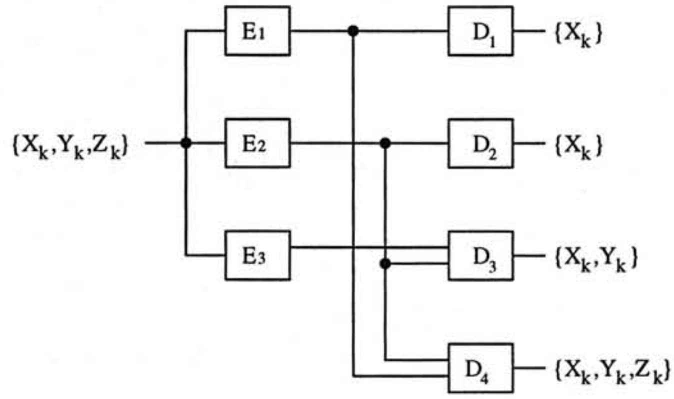


Figure 2.9: MDCS of Example 2.2.6

The equivalent $\mathbf{R_{sp}}$ is given by
$\{(R_1, R_2, R_3):\ \ R_i \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 + R_3 &\geq H(X) + H(Y) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \} \end{aligned}$$

Example 2.2.6 (case 48)
The coding rate region $\mathbf{r_{sp}}$ is given by
$\{(R_1, R_2, R_3):\ \ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} r_1^x &\geq H(X) \\ r_2^x &\geq H(X) \\ r_1^y + r_2^y &\geq H(Y) \\ r_2^y + r_3^y &\geq H(Y) \\ r_1^z + r_2^z &\geq H(Z) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by
$\{(R_1, R_2, R_3):\ \ R_i \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 &\geq H(X) \\ R_2 + R_3 &\geq H(X) + H(Y) \\ R_1 + R_2 &\geq 2H(X) + H(Y) + H(Z) \qquad \} \end{aligned}$$

Example 2.2.7 (case 50)
The coding rate region $\mathbf{r_{sp}}$ is given by

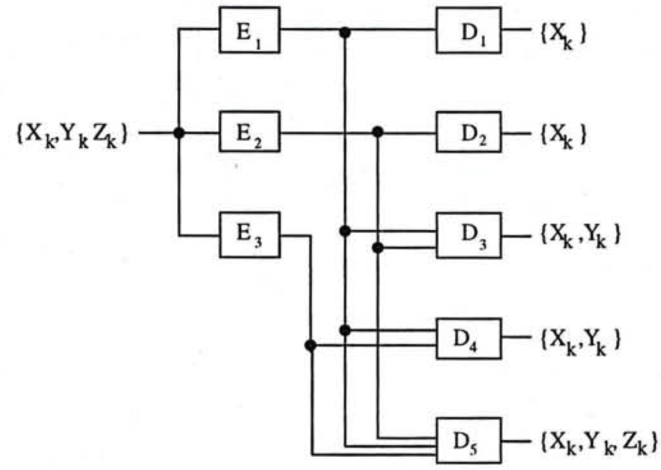


Figure 2.10: MDCS of Example 2.2.7

$\{(R_1, R_2, R_3): \quad R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} r_1^x &\geq H(X) \\ r_2^x &\geq H(X) \\ r_1^y + r_2^y &\geq H(Y) \\ r_1^y + r_3^3 &\geq H(Y) \\ r_1^z + r_2^z + r_3^z &\geq H(Z) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by $\{(R_1, R_2, R_3): \quad R_i \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 &\geq H(X) \\ R_1 + R_2 &\geq H(X) + H(Y) \\ R_2 + R_3 &\geq H(X) + H(Y) \\ R_1 + 2R_2 + R_3 &\geq 2H(X) + 2H(Y) + H(Z) \\ R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \} \end{aligned}$$

Example 2.2.8 (case 63)
The coding rate region $\mathbf{r_{sp}}$ is given by
$\{(R_1, R_2, R_3): R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3$ and

$$\begin{aligned} r_1^x + r_2^x &\geq H(X) \\ r_2^x + r_3^x &\geq H(X) \\ r_1^y + r_3^y &\geq H(Y) \\ r_1^z + r_2^z + r_3^z &\geq H(Z) \qquad \} \end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by

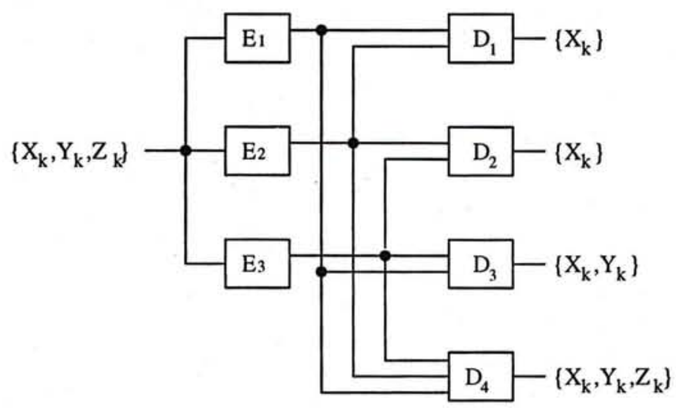


Figure 2.11: MDCS of Example 2.2.8

$\{(R_1, R_2, R_3) : \; R_i \geq 0 \;$ for $i = 1, 2, 3 \;$ and

$$
\begin{aligned}
R_1 + R_2 &\geq H(X) \\
R_2 + R_3 &\geq H(X) \\
R_1 + R_3 &\geq H(X) + H(Y) \\
R_1 + R_2 + R_3 &\geq \frac{3}{2}H(X) + H(Y) + H(Z) \\
2R_1 + R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \\
R_1 + 2R_2 + R_3 &\geq 2H(X) + H(Y) + H(Z) \\
R_1 + R_2 + 2R_3 &\geq 2H(X) + H(Y) + H(Z) \qquad \}
\end{aligned}
$$

In the following three cases in 3-level and 4-level 4-encoder MDCS's for which superposition is optimal are given as a comparison with the 3-encoder cases. It is seen that the complexity of the coding rate region for the 4-encoder MDCS's increases significantly.

Example 2.2.9

The coding rate region $\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, R_3, R_4) : \; R_i = r_i^w + r_i^x + r_i^y + r_i^z$ where $r_i^w, r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3, 4 \;$ and

$$
\begin{aligned}
r_i^w &\geq H(W) \quad 1 \leq i \leq 4 \\
r_1^x + r_3^x &\geq H(X) \\
r_2^x + r_3^x &\geq H(X) \\
r_2^x + r_4^x &\geq H(X) \\
r_1^y + r_2^y + r_3^y &\geq H(Y)
\end{aligned}
$$

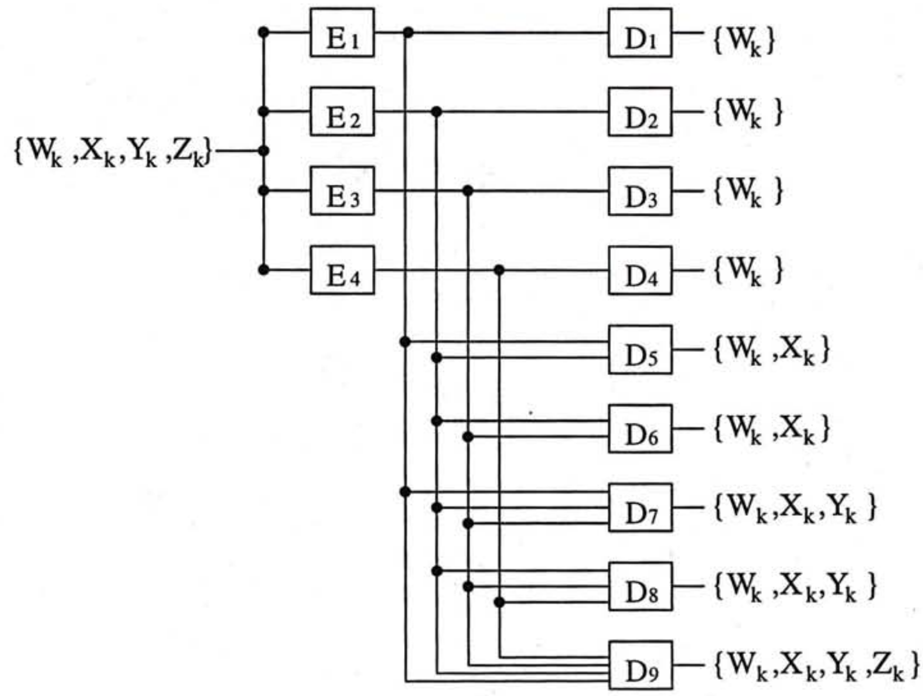


Figure 2.12: MDCS of Example 2.2.9

$$\begin{array}{rcl} r_2^y + r_3^y + r_4^y & \geq & H(Y) \\ r_2^z + r_4^z + r_3^z + r_4^z & \geq & H(Z) \qquad \} \end{array}$$

The equivalent $\mathbf{R_{sp}}$ is given by

$\{(R_1, R_2, R_3, R_4) : \ R_i \geq 0 \ \text{ for } i = 1,2,3,4 \ \text{ and}$

$$\begin{array}{rcl} R_i & \geq & H(W) \ \ 1 \leq i \leq 4 \\ R_1 + R_3 & \geq & 2H(W) + H(X) \\ R_2 + R_3 & \geq & 2H(W) + H(X) \\ R_2 + R_4 & \geq & 2H(W) + H(X) \\ R_1 + R_2 + R_3 & \geq & 3H(W) + H(X) + H(Y) \\ R_2 + R_3 + R_4 & \geq & 3H(W) + H(X) + H(Y) \\ R_1 + R_2 + 2R_3 & \geq & 4H(W) + 2H(X) + H(Y) \\ 2R_2 + R_3 + R_4 & \geq & 4H(W) + 2H(X) + H(Y) \\ R_1 + R_2 + R_3 + R_4 & \geq & 4H(W) + 2H(X) + H(Y) + H(Z) \\ R_1 + 2R_2 + 2R_3 + R_4 & \geq & 6H(W) + 3H(X) + 2H(Y) + H(Z) \qquad \} \end{array}$$

Example 2.2.10

The coding rate region $\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, R_3, R_4) : \ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3,4$ and

$$r_1^x \ \geq \ H(X)$$

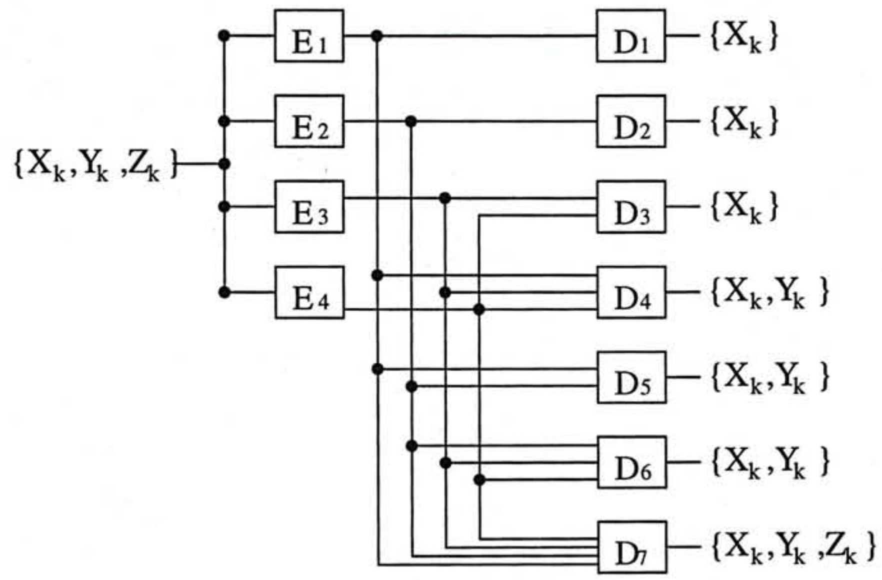


Figure 2.13: MDCS of Example 2.2.10

$$\begin{aligned}
r_2^x &\geq H(X) \\
r_3^x + r_4^x &\geq H(X) \\
r_1^y + r_2^y &\geq H(Y) \\
r_1^y + r_3^y + r_4^y &\geq H(Y) \\
r_3^y + r_3^y + r_4^y &\geq H(Y) \\
r_1^z + r_2^z + r_3^z + r_4^z &\geq H(Z) \qquad \}
\end{aligned}$$

The equivalent $\mathbf{R_{sp}}$ is given by

$\{(R_1, R_2, R_3, R_4) :\ R_i \geq 0$ for $i = 1, 2, 3, 4$ and

$$\begin{aligned}
R_1 &\geq H(X) \\
R_2 &\geq H(X) \\
R_3 + R_4 &\geq H(X) \\
R_1 + R_2 &\geq 2H(X) + H(Y) \\
R_1 + R_3 + R_4 &\geq 2H(X) + H(Y) \\
R_2 + R_3 + R_4 &\geq 2H(X) + H(Y) \\
R_1 + R_2 + R_3 + R_4 &\geq H(X) + \frac{3}{2}H(Y) + H(Z) \\
2R_1 + R_2 + R_3 + R_4 &\geq 4H(X) + 2H(Y) + H(Z) \\
R_1 + 2R_2 + R_3 + R_4 &\geq 4H(X) + H(Y) + H(Z) \\
R_1 + R_2 + 2R_3 + 2R_4 &\geq 4H(X) + 3H(Y) + H(Z) \qquad \}
\end{aligned}$$

Example 2.2.11

The coding rate region $\mathbf{r_{sp}}$ is given by

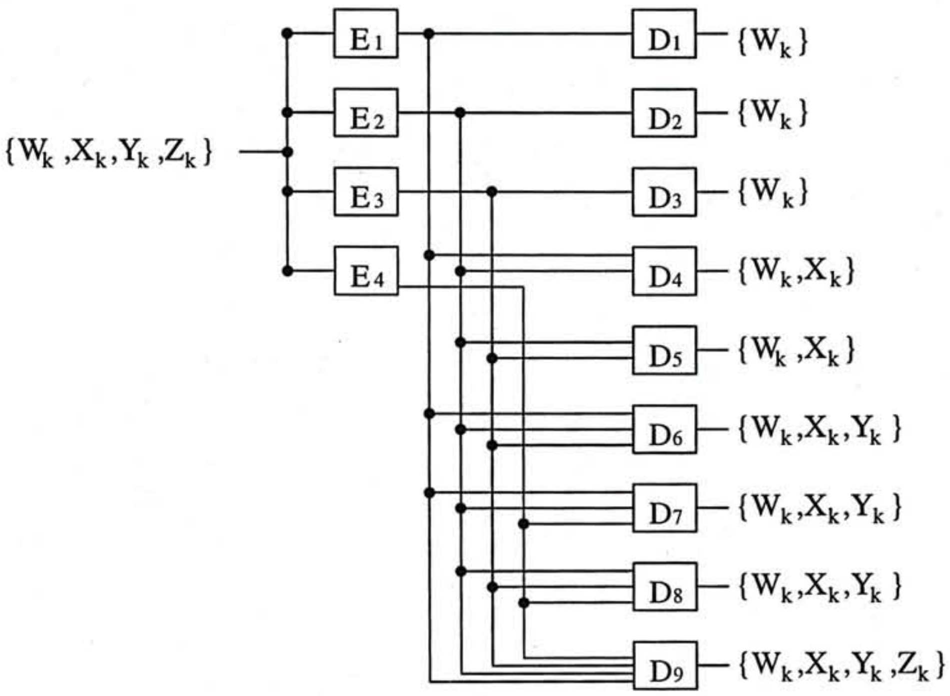


Figure 2.14: MDCS of Example 2.2.11

$\{(R_1, R_2, R_3, R_4): \quad R_i = r_i^w + r_i^x + r_i^y + r_i^z$ where $r_i^w, r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3, 4$ and

$$
\begin{aligned}
r_1^w &\geq H(W) \\
r_2^w &\geq H(W) \\
r_3^w &\geq H(W) \\
r_1^x + r_2^x &\geq H(X) \\
r_2^x + r_3^x &\geq H(X) \\
r_1^x + r_2^x + r_3^y &\geq H(X) \\
r_1^y + r_2^y + r_4^y &\geq H(Y) \\
r_2^y + r_3^y + r_4^y &\geq H(Y) \\
r_1^z + r_2^z + r_3^z + r_4^z &\geq H(Z) \qquad \}
\end{aligned}
$$

The equivalent $\mathbf{R_{SP}}$ is given by

$\{(R_1, R_2, R_3, R_4): \ R_i \geq 0$ for $i = 1, 2, 3, 4$ and

$$
\begin{aligned}
R_1 &\geq H(W) \\
R_2 &\geq H(W) \\
R_3 &\geq H(W) \\
R_1 + R_2 &\geq 2H(W) + H(X) \\
R_2 + R_3 &\geq 2H(W) + H(X) \\
R_2 + R_4 &\geq 2H(W) + H(X)
\end{aligned}
$$

$$
\begin{aligned}
R_1+R_2+R_3 &\geq 3H(W)+H(X)+H(Y)\\
R_1+R_2+R_4 &\geq 2H(W)+H(X)+H(Y)\\
R_2+R_3+R_4 &\geq 2H(W)+H(X)+H(Y)\\
R_1+2R_2+R_3 &\geq 4H(W)+2H(X)+H(Y)\\
2R_2+R_3+R_4 &\geq 3H(W)+2H(X)+H(Y)\\
R_1+R_2+R_3+R_4 &\geq 3H(W)+H(X)+H(Y)+H(Z)\\
2R_1+2R_2+R_3+R_4 &\geq 5H(W)+2H(X)+2H(Y)+H(Z)\\
R_1+2R_2+R_3+2R_4 &\geq 4H(W)+2H(X)+2H(Y)+H(Z)\\
R_1+2R_2+2R_3+R_4 &\geq 4H(W)+2H(X)+2H(Y)+H(Z)\\
R_1+3R_2+R_3+R_4 &\geq 5H(W)+3H(X)+\frac{3}{2}H(Y)+H(Z)\\
2R_1+4R_2+R_3+R_4 &\geq 7H(W)+4H(X)+2H(Y)+H(Z)\\
R_1+4R_2+2R_3+R_4 &\geq 7H(W)+4H(X)+2H(Y)+H(Z)\\
R_1+4R_2+R_3+2R_4 &\geq 6H(W)+4H(X)+2H(Y)+H(Z)\\
2R_1+3R_2+2R_3+2R_4 &\geq 7H(W)+3H(X)+3H(Y)+2H(Z) \quad \}
\end{aligned}
$$

Chapter 3

Symmetrical Multilevel Diversity Coding System

3.1 Introduction

In this chapter we consider a special class of MDCS with symmetrical connectivity between encoders and decoders and we call this class Symmetrical Multilevel Diversity Coding System(SMDCS). In the class of SMDCS, decoders of the same level access the same number of decoders and recover the same number of independent data streams. Moreover, if a level i decoders are defined by accessing k_i encoders in the m-level SMDCS, then any k_i out of m encoders is the fans of a level i decoder. If we treat the decoders that access exactly the same set of encoders as identical decoders, then there are $C^m_{k_i}$ different decoders within level i ($C^i_j = \frac{i!}{(i-j)!j!}$). Also we can implement a maximum of m levels of decoders in such a m-encoder SMDCS. Rate distortion approach to this class of MDCS was considered in Roche *et al*[14]. In Figure 3.1, we show one member of SMDCS with three encoders, three data streams and three levels of decoders.

We discover that superposition is *optimal* for all the cases we have studied in the class of SMDCS. This much simplify the coding of information in the system. The class of SMDCS have significant applications in various situations. In the following we list three situations in which SMDCS may find applicable.

1. In a m-transmitter communication System. We may assume any of the transmitters may break down or any of the channels used by the transmitters may be

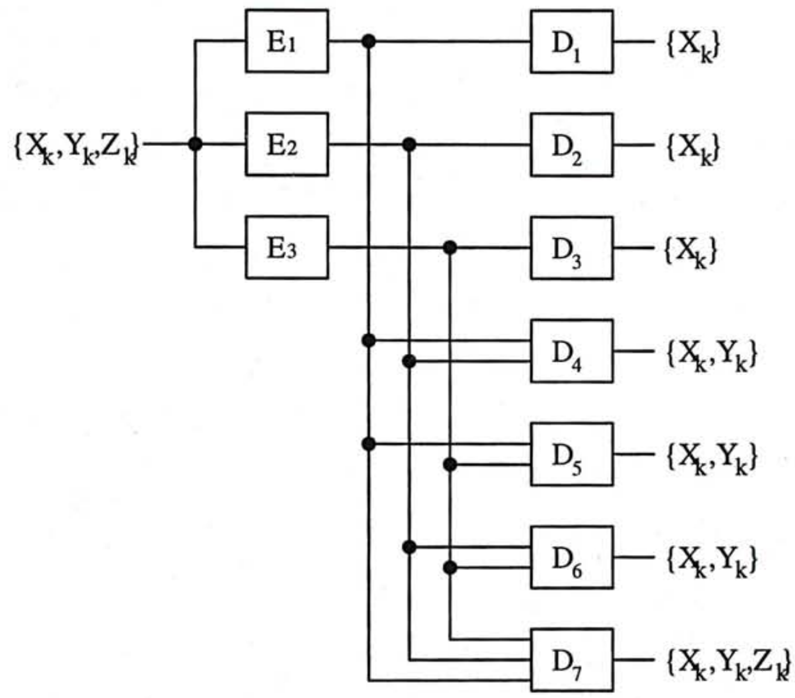


Figure 3.1: Configuration of an SMDCS with 3 encoders and 3 data streams

blocked. In case of some of the transmitters break down or some of the channels blocked, we can access only k transmitters where $k < m$. This subset in general is one of the C_k^m subsets. Under our symmetrical assumption in this class of MDCS, what we can recover from any of the C_k^m subsets is the same degraded version of information, no matter which k transmitters we can access.

2. In a computer network, it is more general to assume packets are route independently to the destination. Suppose that the m packets carries k messages $M_1, M_2, \ldots, M_k$ in decreasing order of importance and $k \leq m$. Information can be recovered from any k packets first arrive. Even if there are i packets arrived after a delay threshold set by the receiver or $k - i$ packets are in error among the k arrived and $i < k$, we can still recover the most important i messages.

3. In a secret sharing system more secret is revealed if more people pool their knowledge. To maintain *fairness* in the system, any k persons out of the n-person-group should be able to reveal secret up to the same level. When the whole group come together, of course a maximum amount of secret is reveal. But even for any single person or two-person-group, some common primary informaion can be obtained.

We use $SMDCS[k, m, (m_1, m_2, ..., m_k)]$ to denote each member of this class of MDCS with an information source consisting of k independent data streams of decreasing importance. Also there are m encoders. Each level-i decoders can access m_i encoders to recover stream 1 to streams i. The corresponding coding rate region induced by superposition is denoted as $\mathbf{R_{kmm_1m_2...m_k}}$. Note that $k \geq m$ and $m_1 < m_2 < \ldots < m_k$. The most complete configuration of a member of this class is $SMDCS[m, m, (1, 2, ..., m)]$ in which one more data stream is recovered by accessing one more encoder. We allow the degeneration of some $m - k$ data streams and some number in $(1, 2, ..., m)$ being removed accordingly so we obtain the more general notation $SMDCS[k, m, (m_1, m_2, ..., m_k)]$.

Since superposition is a feasible coding scheme, in general the coding rate region induced by superposition $\mathbf{R_{sp}}$ is a subset of $\mathbf{R}$, the admissible coding rate region. We will prove the optimality of superposition by showing that $\mathbf{R}$ is a subset of $\mathbf{R_{sp}}$. So $\mathbf{R} = \mathbf{R_{sp}}$. Therefore $\mathbf{R_{sp}}$ is a sufficient characterization of the general admissible coding rate region $\mathbf{R}$. Throughout this chapter we have to use *rate* constraints to prove the converse of the coding scheme by superposition. The method to construct the *rate* constraints from the *subrate* constraints and the proof of equivalence between the two is deferred to Chapter 4 and Appendix A.

Before we study the different subclasses of SMDCS, we first give a proof of a theorem which provide a lower bound on the rate sum of the encoders in a SMDCS. This result is originally obtained in Roche [13]. In the following, the ith data stream is denoted by $\{X_k^i\}$.

Theorem 1 *In a $SMDCS[m, m, (1, 2, ..., m)]$ where the coding rate of encoder i is R_i,*

$$\sum_{k=1}^{m} R_k \geq m \sum_{k=1}^{m} \frac{H(X^k)}{k} \tag{3.1}$$

Proof

Let $E \subseteq \{S_1, S_2, S_3, \ldots, S_m\}$ where S_i is the output random variable of encoder i.

$$\begin{aligned} R_k &\geq H(S_k) \quad \text{for } 1 \leq k \leq m \\ &= n^{-1} H(S_k, (X^1)_1^n) \end{aligned}$$

$$= \quad n^{-1}H(S_k|(X^1)_1^n) + H(X^1) \tag{3.2}$$

So,

$$\sum_{k=1}^{m} R_k \geq mH(X^1) + \frac{m}{nC_1^m} \sum_{E:|E|=1} H(E|(X^1)_1^n) \tag{3.3}$$

Moreover,

$$\begin{aligned}
H(X^k) &= n^{-1}H((X^k)_1^n) &(3.4)\\
&= n^{-1}H((X^k)_1^n|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n) &(3.5)\\
&= n^{-1}H((X^k)_1^n|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n,E:|E|=k) &(3.6)\\
&\quad +n^{-1}I((X^k)_1^n;E:|E|=k|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n)\\
&= n^{-1}I((X^k)_1^n;E:|E|=k|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n) &(3.7)
\end{aligned}$$

Since (3.7) is true for all E such that $|E| = k$, so

$$H(X^k) = \frac{1}{nC_k^m} \sum_{E:|E|=k} I(E;(X^k)_1^n|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n)$$

$$\frac{mH(X^k)}{k} - \frac{m}{nC_k^m} \sum_{E:|E|=k} \frac{I(E;(X^k)_1^n|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n)}{k} = 0 \tag{3.8}$$

By summing up (3.8) for $2 \leq k \leq m$, we have

$$\sum_{k=2}^{m}[\frac{mH(X^k)}{k} - \frac{m}{nC_k^m} \sum_{E:|E|=k} \frac{I(E;(X^k)_1^n|(X^{k-1})_1^n,\ldots,(X^1)_1^n)}{k}] = 0 \tag{3.9}$$

Now, by (3.3)

$$\begin{aligned}
\sum_{k=1}^{m} R_k &\geq mH(X^1) + \frac{m}{nC_1^m} \sum_{E:|E|=1} H(E|(X^1)_1^n)\\
+\sum_{k=2}^{m}[\frac{mH(X^k)}{k} &- \frac{m}{nC_k^m} \sum_{E:|E|=k} \frac{I(E;(X^k)_1^n|(X^{k-1})_1^n,\ldots,(X^1)_1^n)}{k}] \qquad (3.10)\\
&= m\sum_{k=1}^{m} \frac{H(X^k)}{k} + \frac{m}{nC_1^m} \sum_{E:|E|=1} H(E|(X^1)_1^n)\\
&+ \frac{m}{n}\{\sum_{k=2}^{m}[\frac{1}{C_k^m} \sum_{E:|E|=k} \frac{H(E|(X^k)_1^n,(X^{k-1})_1^n,\ldots,(X^1)_1^n)}{k}]\\
&- \sum_{k=2}^{m}[\frac{1}{C_k^m} \sum_{E:|E|=k} \frac{H(E|(X^{k-1})_1^n,(X^{k-2})_1^n,\ldots,(X^1)_1^n)}{k}]\}\\
&= m\sum_{k=1}^{m} \frac{H(X^k)}{k}
\end{aligned}$$

$$+ \quad \frac{m}{n}\{\sum_{k=1}^{m}[\frac{1}{C_k^m}\sum_{E:|E|=k}\frac{H(E|(X^k)_1^n,(X^{k-1})_1^n,\ldots,(X^1)_1^n)}{k}]$$
$$- \quad \sum_{k=1}^{m-1}[\frac{1}{C_{k+1}^m}\sum_{E:|E|=k+1}\frac{H(E|(X^k)_1^n,\ldots,(X^1)_1^n)}{k+1}]\} \tag{3.11}$$

(3.2) is true since $(X^1)_1^n$ is a function of S_i for $1 \le i \le m$. Also (3.7) is true since $(X^k)_1^n$ is a function of $(S_{i_1}, S_{i_2}, \ldots, S_{i_k})$ for $1 \le i_1 < \ldots < i_k \le m$. (3.5) is true because the data streams are independent. The inequality sign of (3.2) is preserved in (3.10) because the second term of (3.10) is just zero by (3.9). Here we introduce the Han's Inequality [5], which shows the average entropy per elements of a subset of a set of random variables $\{X^i, 1 \le i \le m\}$ decreases as the size of the subset increases. Let

$$H_k^m = \frac{1}{C_k^m}\sum_{E:|E|=k}\frac{H(E)}{k}$$

Han's Inequality states that

$$H_1^m \ge H_2^m \ge \ldots \ge H_m^m$$

We see that we can apply the Han's inequality (conditioned on $((X^k)_1^n, (X^{k-1})_1^n, \ldots, (X^1)_1^n)$ for $1 \le k \le m-1$) on the the last two terms in (3.11) and conclude that the difference is always greater than zero. So (3.11) becomes

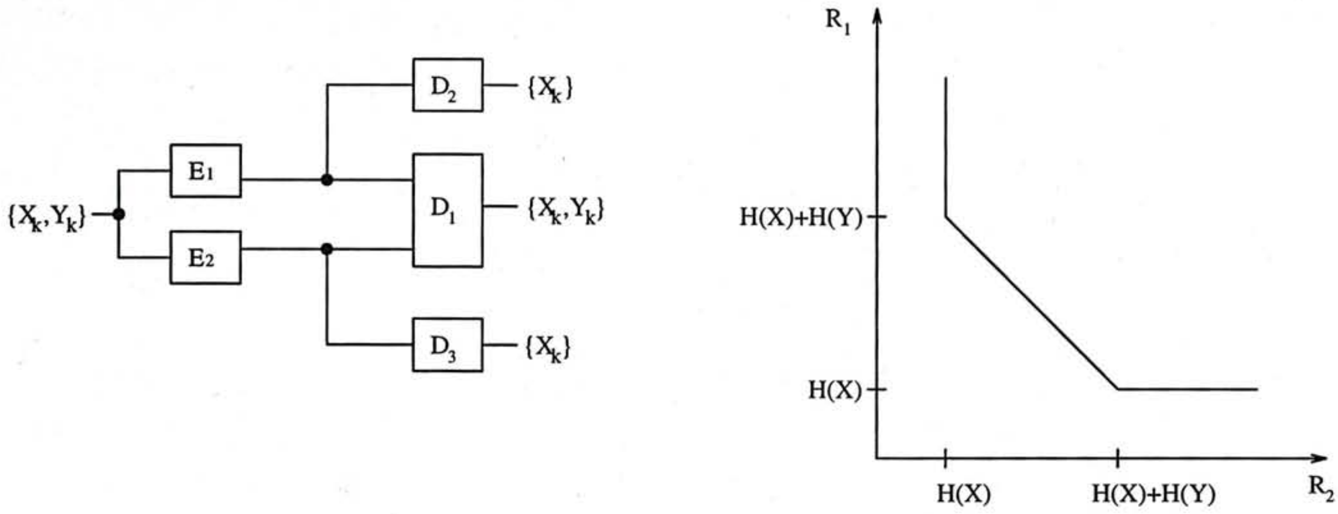
$$\sum_{k=1}^{m} R_k \ \ge \ m\sum_{k=1}^{m}\frac{H(X^k)}{k} + \frac{H(E:|E|=m|(X^m)_1^n,(X^{m-1})_1^n,\ldots,(X^1)_1^n)}{n}$$
$$\ge \ m\sum_{k=1}^{m}\frac{H(X^k)}{k} \tag{3.12}$$

So the theorem holds. □

We will apply this theorem in the proof of the optimality of superposition.

3.2 SMDCS[2,m,(1,m)]

In this section we consider a subclass of SMDCS with m encoders and 2 independent data streams and decoders of 2 levels. The simlpest case in this class $SMDCS[2,2,(1,2)]$ which is shown in Figure 3.2. Assume the 2 data streams are represented by $\{X_k\}$ and $\{Y_k\}$.



SMDCS[2,2,(1,2)] Configuration SMDCS[2,2,(1,2)] Coding Rate Region

Figure 3.2: SMDCS[2,2,(1,2)]

The coding rate region induced by superposition in terms of subrate constraints $\mathbf{r_{2212}}$ is given by

$\{ (R_1, R_2):$

$$\begin{aligned} R_i &= r_i^x + r_i^y \quad i = 1,2 \\ \text{and} & \\ r_1^x &\geq H(X) \\ r_2^x &\geq H(X) \\ r_1^y + r_2^y &\geq H(Y) \qquad \} \end{aligned}$$

An equivalent coding rate region in terms of rate constraints, denoted as $\mathbf{R_{2212}}$, is given by

$\{ (R_1, R_2):$

$$\begin{aligned} R_1 &\geq H(X) \\ R_2 &\geq H(X) \\ R_1 + R_2 &\geq 2H(X) + H(Y) \qquad \} \end{aligned}$$

$\mathbf{R_{2212}}$ is shown in Figure 3.2.

Theorem 2 *In a $SMDCS[2,2,(1,2)]$, the admissible coding rate region* $\mathbf{R}$ *is given by* $\mathbf{R_{2212}}$ *, and superposition is optimal.*

Proof

Superposition is a feasible coding scheme so $\mathbf{R_{2212}} \subseteq \mathbf{R}$

For any admissible tuple (R_1, R_2),

$R_1 \geq n^{-1}H(S_1) = n^{-1}H(S_1, X_1^n) = n^{-1}H(S_1|X_1^n) + H(X) \geq H(X)$

First equality is true because (X_1^n) is a function of S_1. Similarly (3.13) is true. We see that $SMDCS[2,2,(1,2)]$ is a special case of the SMDCS described in Theorem (1). (3.13) can be obtained by involking Theorem (1). So $\mathbf{R} = \mathbf{R_{sp}}$ and so superposition is the optimal coding scheme. □

We can generalize $SMDCS[2,2,(1,2)]$ to $SMDCS[2,m,(1,m)]$. Note that $SMDCS$ $[2,m,(1,m)]$ is obtained by degenerating $\{X_k^i\}$ for $2 \leq i \leq m-1$ in $SMDCS[m,m,(1,2,...,m)]$.

The coding rate region induced by superposition $\mathbf{r_{2m1m}}$ in terms of subrate constraints is given by

$\{\ (R_1, R_2, \ldots, R_m):$

$$R_i = r_i^x + r_i^y \quad \text{for } i = 1, 2, \ldots, m$$

and

$$\begin{aligned} r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq m \\ r_1^y + r_2^y + \ldots + r_m^y &\geq H(Y) \qquad \} \end{aligned}$$

An equivalent description denoted as $\mathbf{R_{2m1m}}$ is given by

$\{(R_1, R_2, \ldots, R_m):$

$$\begin{aligned} R_i &\geq H(X) \quad \text{for } 1 \leq i \leq m \\ R_1 + R_2 + \ldots + R_m &\geq mH(X) + H(Y) \end{aligned}$$

Corollary 1 *The admissible coding rate region for $SMDCS[2,m,(1,m)]$ is given by* $\mathbf{R_{2m1m}}$ *and superposition is optimal.*

Proof

The equivalence of $\mathbf{r_{2m1m}}$ and $\mathbf{R_{2m1m}}$ is proved in Appendix A. The forward part is true since superposition is feasible. For any admissible m-tuple $(R_1, R_2, \ldots, R_m)$ it is easily seen that (3.13) is satisfied by admissibility condition. (3.13) is satisfied by invoking Theorem (1) and setting $X^1 = X, (X^m) = Y$ and $X^i = 0$ for $2 \leq i \leq m-1$. □

At the moment we cannot solve explicitly for the general case of $SMDCS[2, m, (1, k)]$ where $m \geq k$. We know that $\mathbf{r_{2m1k}}$ is given by

$\{(R_1, R_2, \ldots, R_k):$

$$\begin{aligned} r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq m \\ r_{i_1}^y + r_{i_2}^y + \ldots + r_{i_k}^y &\geq H(Y) \quad \text{for } 1 < i_1 \leq \ldots < i_k \leq m \qquad \} \end{aligned}$$

We *conjecture* that $\mathbf{R_{2m1k}}$ is given by

$\{(R_1, R_2, \ldots, R_k):$

$$R_i \geq H(X) \quad \text{for } 1 \leq i \leq m \tag{3.13}$$

$$R_{i_1} + R_{i_2} + \ldots + R_{i_k} \geq kH(X) + H(Y) \quad \text{for } 1 \leq i_1 < \ldots < i_k \leq m \tag{3.14}$$

The converse can be easily proved for $\mathbf{R_{2m1k}}$. We can solve for more complicated subclasses of SMDCS in later sections and thoses SMDCS's can be reduced to some cases in $SMDCS[2, m, (1, k)]$ and the conjecture is true for those cases.

A special case of $SMDCS[2, m, (1, k)]$, $SMDCS[2, 4, (1, 3)]$ is shown in Figure 3.3.

3.3 SMDCS[3,m,(1,2,m)]

In this section we consider a subclass of SMDCS in which each system has m encoders, 3 data streams and the decoders are divided into 3 levels. First we consider the case $SMDCS[3, 3, (1, 2, 3)]$, which is the simplest in this subclass. The configuration of $SMDCS[3, 3, (1, 2, 3)]$ is shown in Figure 3.1 at the begining of this chapter as an example of SMDCS. Assume the 3 independent data streams are represented as $\{X_k\}, \{Y_k\}$ and $\{Z_k\}$.

Also, every two encoders in $SMDCS[3, 3, (1, 2, 3)]$ and the decoders which include only the two encoders as their fan forms a $SMDCS[2, 2, (1, 2)]$.

The coding rate region described by subrate constraints, denoted as $\mathbf{r_{33123}}$ is

$\{(R_1, R_2, R_3):$

$$R_i = r_i^x + r_i^y + r_i^z \quad \text{for } 1 \leq i \leq 3 \tag{3.15}$$

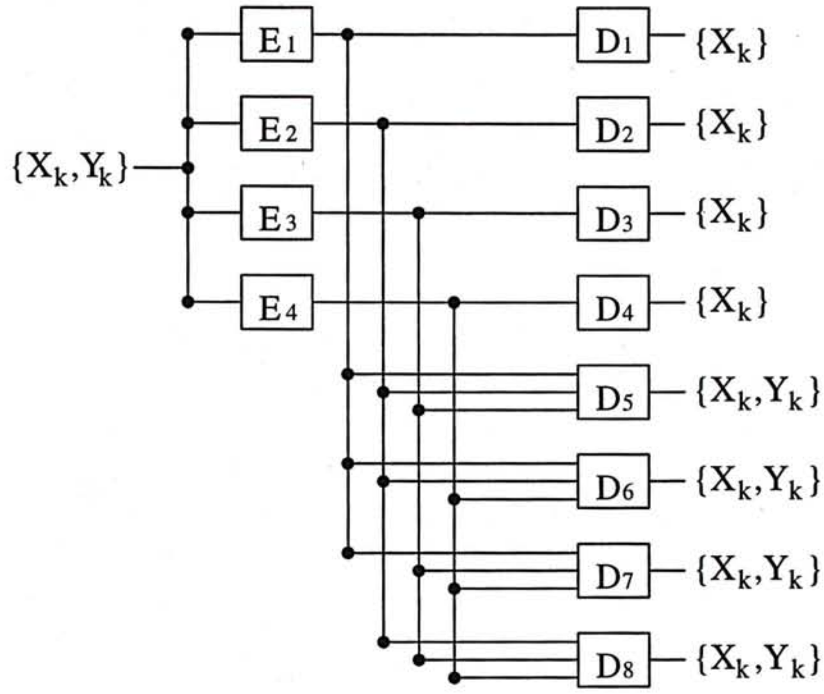


Figure 3.3: Configuration of SMDCS[2,4,(1,3)]

and

$$\begin{aligned} r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq 3 & (3.16) \\ r_i^y + r_j^y &\geq H(Y) \quad \text{for } 1 \leq i < j \leq 3 & (3.17) \\ r_1^z + r_2^z + r_3^z &\geq H(Z) \quad \} & (3.18) \end{aligned}$$

The coding rate region described by rate constraints, denoted by $\mathbf{R_{33123}}$, is given by
$\{(R_1, R_2, R_3):$

$$\begin{aligned} R_i &\geq H(X) \quad \text{for } 1 \leq i \leq 3 & (3.19) \\ R_i + R_j &\geq 2H(X) + H(Y) \quad \text{for } 1 \leq i < j \leq 3 & (3.20) \\ R_1 + R_2 + R_3 &\geq 3H(X) + \frac{3}{2}H(Y) + H(Z) & (3.21) \\ 2R_i + R_{i\oplus 1} + R_{i\oplus 2} &\geq 4H(X) + 2H(Y) + H(Z) \quad \} & (3.22) \end{aligned}$$

where $\oplus$ is defined by

$$x \oplus y = \begin{cases} x + y & \text{if } x + y \leq 3 \\ x + y - 3 & \text{if } x + y > 3 \end{cases}$$

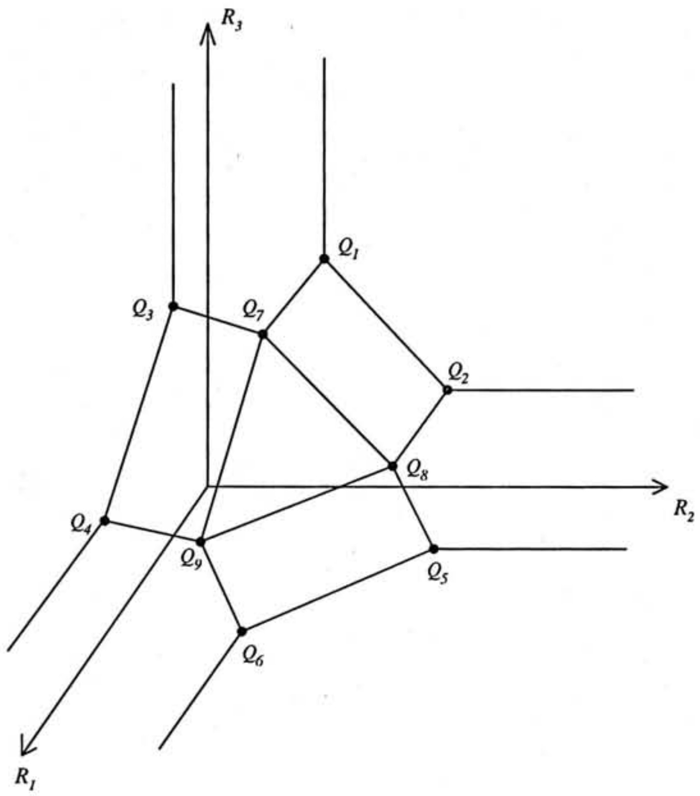


Figure 3.4: Coding Rate Region of SMDCS[3,3(1,2,3)]

$\mathbf{R_{33123}}$ is shown in Figure 3.4. Now we prove the above coding rate region is in fact the general admissible coding rate region.

Theorem 3 *For a $SMDCS[3,3,(1,2,3)]$, the admissible coding rate region is given by $\mathbf{R_{33123}}$ and superposition is the optimal coding scheme.*

Proof

The equivalence of $\mathbf{r_{33123}}$ and $\mathbf{R_{33123}}$ will be proved in Appendix A. Superposition is a feasible coding scheme so $\mathbf{R_{33123}}$ is a subset of the admissible coding rate region. Now for any admissible 3-tuple (R_1, R_2, R_3),

$$R_i \geq n^{-1}H(S_i) = n^{-1}H(S_i, X_1^n) \geq H(X) + n^{-1}H(S_i|X_1^n) \geq H(X) \quad \text{for } 1 \leq i \leq 3$$

We also see that $SMDCS[3,3,(1,2,3)]$ is a special case of the SMDCS in Theorem (1). Also we also observe that every two encoders form a SMDCS[2,2,(1,2)]. So (3.20) and (3.21) can be proved by Theorem (1). Also,

$$2R_1 + R_2 + R_3 \;\; \geq \;\; n^{-1}(2H(S_1) + H(S_2) + H(S_3))$$

$$
\begin{aligned}
&= n^{-1}(2H(S_1|X_1^n) + H(S_2|X_1^n) + H(S_3|X_1^n)) + 4H(X) \\
&\geq n^{-1}(H(S_1, S_2, Y_1^n|X_1^n) + H(S_1, S_3, Y_1^n|X_1^n)) + 4H(X) \\
&= n^{-1}(H(S_1, S_2|X_1^n, Y_1^n) + H(S_1, S_3|Y_1^n, X_1^n)) + 4H(X) + 2H(Y) \\
&\geq n^{-1}(H(S_1, S_2, S_3, Z_1^n|X_1^n, Y_1^n)) + 4H(X) + 2H(Y) \\
&\geq 4H(X) + 2H(Y) + H(Z)
\end{aligned}
$$

First equalities holds because (X_1^n) is a function of S_i for $1 \leq i \leq 3$. Second inequality holds because (Y_1^n) is a function of (S_i, S_j) for $1 \leq i < j \leq 3$. Third inequlity holds because (Z_1^n) is a function of (S_1, S_2, S_3). Other two constraints in (3.22) is proved similarly since they have symmetrical structure.

So $\mathbf{R} = \mathbf{R_{33123}}$. Superposition is the optimal coding scheme thereafter. □

Here we gives a few remarks about $SMDCS[3, 3, (1, 2, 3)]$ and see how it is related with other cases of SMDCS.

1. $SMDCS[2, 3, (2, 3)]$ is a special case of $SMDCS[3, 3, (1, 2, 3)]$ by degenerate stream $\{X_k\}$. The admissible coding rate region is obtained by setting $H(X) = 0$ in the constraints of $\mathbf{R_{33123}}$. We don't need to delete any redundant constraints from $\mathbf{R_{33123}}$ in this special case.

2. $SMDCS[2, 3, (1, 2)]$ and $SMDCS[2, 3, (1, 3)]$ are special case of $SMDCS[2, m, (1, k)]$. They are also special cases of $SMDCS[3, 3, (1, 2, 3)]$. In this case $\mathbf{R}$ of $SMDCS$ $[2, 3, (1, 2)]$ is obtained by setting $H(Z) = 0$ in the constraints. But then (3.21) and (3.22) are implied by (3.20) so are redundant and deleted for a minimal description of $\mathbf{R}$. Also $\mathbf{R}$ of $SMDCS[2, 3, (1, 3)]$ are obtained by setting $H(Y) = 0$ in the constraints. (3.20) and (3.22) are redundant and should be deleted. The result is consistent with our conjecture in the last section.

We *conjecture* that the admissible coding rate region of $SMDCS[3, m, (1, 2, 3)]$ is given by

$\{(R_1, R_2, \ldots, R_m):$

$$
\begin{aligned}
R_i &\geq H(X) \quad \text{for } 1 \leq i \leq m \\
R_i + R_j &\geq 2H(X) + H(Y) \quad \text{for } 1 \leq i < j \leq m
\end{aligned}
$$

$$\begin{aligned} R_i + R_j + R_k &\geq 3H(X) + \frac{3}{2}H(Y) + H(Z) \\ 2R_i + R_j + R_k &\geq 4H(X) + 2H(Y) + H(Z) \\ R_i + 2R_j + R_k &\geq 4H(X) + 2H(Y) + H(Z) \\ R_i + R_j + 2R_k &\geq 4H(X) + 2H(Y) + H(Z) \quad \text{for } 1 \leq i < j < k \leq m \quad \} \end{aligned}$$

And superposition is the optimal coding scheme.

We are unable to prove this explicitly but we can verify that it is true for the case $SMDCS[3,4,(1,2,3)]$ which can be obtained from the result of $SMDCS[4,4,(1,2,3,4)]$ in the last section of this chapter.

$SMDCS[3,3,(1,2,3)]$ also belongs to the general subclass $SMDCS[3,m,(1,2,m)]$. We are able to solve two more cases in this subclass which are $SMDCS[3,4,(1,2,4)]$ and $SMDCS[3,5,(1,2,5)]$. First we show the case of $SMDCS[3,4,(1,2,4)]$.

The coding rate region described by subrate constraints denoted by $\mathbf{r_{34124}}$ is $\{(R_1, R_2, R_3, R_4):$

$$\begin{aligned} R_i &= r_i^x + r_i^y + r_i^z \quad \text{for } 1 \leq i \leq 4 \\ \text{and} \\ r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq 4 \\ r_i^y + r_j^y &\geq H(Y) \quad \text{for } 1 \leq i < j \leq 4 \\ r_1^z + r_2^z + r_3^z + r_4^z &\geq H(Z) \quad \} \end{aligned}$$

The coding rate region described by rate constraints, denoted as $\mathbf{R_{34124}}$, is given by
$\{(R_1, R_2, R_3, R_4):$

$$\begin{aligned} R_i &\geq H(X) \quad \text{for } 1 \leq i \leq 4 \\ R_i + R_j &\geq 2H(X) + H(Y) \quad \text{for } 1 \leq i < j \leq 4 \\ R_1 + R_2 + R_3 + R_4 &\geq 3H(X) + 2H(Y) + H(Z) \\ 3R_i + R_{i\oplus 1} + R_{i\oplus 2} + R_{i\oplus 3} &\geq 6H(X) + 3H(Y) + H(Z) \quad \text{for } 1 \leq i \leq 4 \quad \} \end{aligned}$$

here $\oplus$ is defined by

$$x \oplus y = \begin{cases} x+y & \text{if } x+y \leq 4 \\ x+y-4 & \text{if } x+y > 4 \end{cases}$$

Next we show the case of $SMDCS[3,5,(1,2,5)]$.

The coding rate region described by subrate constraints $\mathbf{r_{35125}}$ is

$\{(R_1, R_2, R_3, R_4, R_5):$

$$\begin{aligned} R_i &= r_i^x + r_i^y + r_i^z \quad \text{for } 1 \leq i \leq 5 \\ &\text{and} \\ r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq 5 \\ r_i^y + r_j^y &\geq H(Y) \quad \text{for } 1 \leq i < j \leq 5 \\ r_1^z + r_2^z + r_3^z + r_4^z + r_5^z &\geq H(Z) \quad \} \end{aligned}$$

The coding rate region described by rate constraints, denoted as $\mathbf{R_{35125}}$, is given by

$\{(R_1, R_2, R_3, R_4, R_5):$

$$\begin{aligned} R_i &\geq H(X) \quad \text{for } 1 \leq i \leq 5 \\ R_i + R_j &\geq 2H(X) + H(Y) \quad \text{for } 1 \leq i < j \leq 5 \\ R_1 + R_2 + R_3 + R_4 + R_5 &\geq 5H(X) + \frac{5}{2}H(Y) + H(Z) \\ 4R_i + R_{i\oplus 1} + R_{i\oplus 2} + R_{i\oplus 3} + R_{i\oplus 4} &\geq 8H(X) + 4H(Y) + H(Z) \text{ for } 1 \leq i \leq 5 \quad \} \end{aligned}$$

here $\oplus$ is defined by

$$x \oplus y = \begin{cases} x+y & \text{if } x+y \leq 5 \\ x+y-5 & \text{if } x+y > 5 \end{cases}$$

We conjectured that $SMDCS[3,m,(1,2,m)]$ have the following coding rate region but we are unable to prove it explicitly.

$\{(R_1, R_2, \ldots, R_m):$

$$R_i \geq H(X) \quad \text{for } 1 \leq i \leq m$$

$$\begin{aligned}
R_i + R_j &\geq 2H(X) + H(Y) \quad \text{for } 1 \leq i < j \leq m \\
R_1 + R_2 + \ldots + R_m &\geq mH(X) + \frac{m}{2}H(Y) + H(Z) \\
(m-1)R_i + R_{i\oplus 1} + R_{i\oplus 2} + \ldots + R_{i\oplus m} &\geq 2(m-1)H(X) + (m-1)H(Y) + H(Z) \\
&\quad \text{for } 1 \leq i \leq m \quad \}
\end{aligned}$$

here $\oplus$ is defined by

$$x \oplus y = \begin{cases} x + y & \text{if } x + y \leq m \\ x + y - m & \text{if } x + y > m \end{cases}$$

Remark

Some subtle correlation exist between the contraints and the extreme points of the coding rate regions of $SMDCS[3,3,(1,2,3)]$, $SMDCS[3,4,(1,2,5)]$ and $SMDCS$ $[3,5,(1,2,5)]$ though we cannot derive any logical relation between them. In the following we list the set of extreme points for the 3 SMDCS's. We can check that all the extreme points of $\mathbf{R_{sp}}$ are in $\mathbf{r_{sp}}$. We just list the typical ones from the whole set of extreme points, the rest are obtained by permutate the order of the coordinates of the points in the list since they have symmetrical stucture. (In the following $\times 3$ means there are 3 similar cases symmetrical to one another).

$SMDCS$	extreme point of coding rate region
$SMDCS[3,3,(1,2,3)]$	$(H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}H + H(Z))$ $(\times 3)$
	$(H(X), H(X) + H(Y), H(X) + H(Y) + HZ))$ $(\times 6)$
$SMDCS[3,4,(1,2,4)]$	$(H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2} + H(Z))$ $(\times 4)$
	$(H(X), H(X) + H(Y), H(X) + H(Y), H(X) + H(Y) + HZ))$ $(\times 12)$
$SMDCS[3,5,(1,2,5)]$	$(H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2} + H(Z))$ $(\times 5)$
	$(H(X), H(X) + H(Y), H(X) + H(Y), H(X) + HY), H(X) + H(Y) + H(Z))$ $(\times 20)$

3.4 SMDCS[3,m,(1,3,m)]

At the moment we are able to solve 2 cases in this subclass which are $SMDCS$ $[3,4,(1,3,4)]$ and $SMDCS[3,5,(1,3,5)]$. $SMDCS[3,4,(1,3,4)]$ are shown in the Figure 3.5.

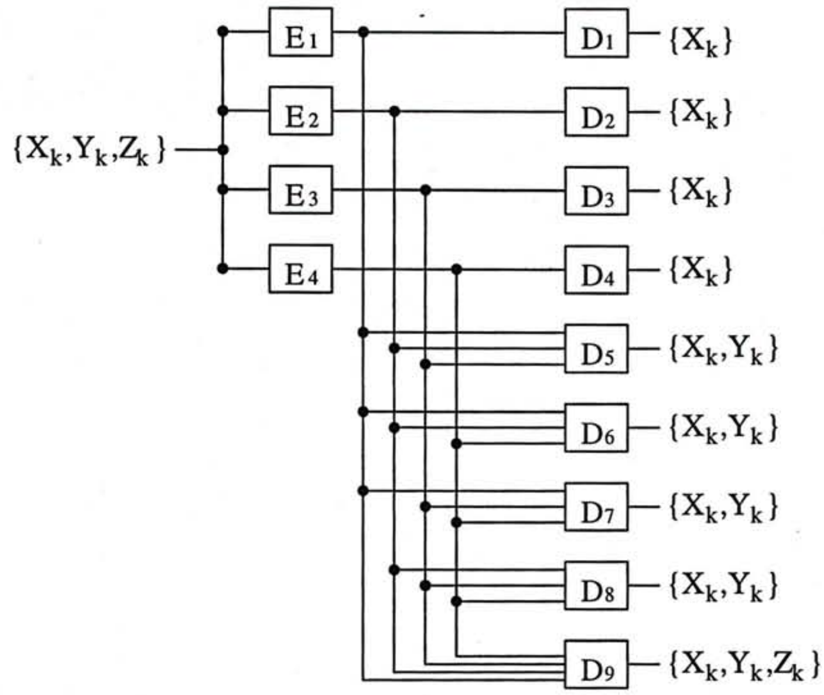


Figure 3.5: Configuration of SMDCS[3,4,(1,3,4)]

The coding rate region of $SMDCS[3,4,(1,3,4)]$ induced by superposition denoted by $\mathbf{r_{34134}}$ is given by

$\{(R_1, R_2, R_3, R_4):$

$$R_i = r_i^x + r_i^y + r_i^z \quad \text{for } 1 \le i \le 4$$

and

$$\begin{aligned} r_i^x &\ge H(X) \quad \text{for } 1 \le i \le 4 \\ r_i^y + r_j^y + r_k^y &\ge H(Y) \quad \text{for } 1 \le i < j < k \le 4 \\ r_1^z + r_2^z + r_3^z + r_4^z &\ge H(Z) \quad \} \end{aligned}$$

Define $\pi(1234)$ as the set of all permutations of the sequence (1,2,3,4). The coding rate region described by rate constraints, denoted as $\mathbf{R_{34134}}$, is given by

$\{(R_1, R_2, R_3, R_4):$

$$\begin{aligned} R_i &\ge H(X) \quad \text{for } 1 \le i \le 4 \\ R_i + R_j + R_k &\ge 3H(X) + H(Y) \\ R_1 + R_2 + R_3 + R_4 &\ge 4H(X) + \frac{4}{3}H(Y) + H(Z) \end{aligned}$$

$$
\begin{aligned}
3R_i + 2R_j + 2R_k + 2R_l &\geq 9H(X) + 3H(Y) + 2H(Z) \\
2R_i + 2R_j + R_k + R_l &\geq 6H(X) + 2H(Y) + H(Z) \\
&\quad \forall (i,j,k,l) \in \pi(1,2,3,4) \quad \}
\end{aligned}
$$

The coding rate region $\mathbf{r_{35135}}$ of $SMDCS[3,5,(1,3,5)]$ in terms of subrate constraints is given by

$\{(R_1, R_2, R_3, R_4, R_5):$

$$
R_i = r_i^x + r_i^y + r_i^z \quad \text{for } 1 \leq i \leq 4
$$

and

$$
\begin{aligned}
r_i^x &\geq H(X) \quad \text{for } 1 \leq i \leq 4 \\
r_i^y + r_j^y + r_k^y &\geq H(Y) \quad \text{for } 1 \leq i < j < k \leq 4 \\
r_1^z + r_2^z + r_3^z + r_4^z + r_5^z &\geq H(Z) \quad \}
\end{aligned}
$$

Define $\pi(12345)$ as the set of all permutation of the sequence (1,2,3,4,5). The coding rate region described by rate constraints denoted by $\mathbf{R_{35135}}$ is given by

$\{(R_1, R_2, R_3, R_4, R_5):$

$$
\begin{aligned}
R_i &\geq H(X) \quad \text{for } 1 \leq i \leq 4 && (3.23) \\
R_i + R_j + R_k &\geq 3H(X) + H(Y) && (3.24) \\
R_1 + R_2 + R_3 + R_4 + R_5 &\geq 5H(X) + \frac{5}{3}H(Y) + H(Z) && (3.25) \\
2R_i + R_j + R_k + R_l + R_m &\geq 6H(X) + 2H(Y) + H(Z) && (3.26) \\
3R_i + 3R_j + R_k + R_l + R_m &\geq 9H(X) + 3H(Y) + H(Z) && (3.27) \\
&\quad \forall (i,j,k,l,m) \in \pi(1,2,3,4,5) \quad \} && (3.28)
\end{aligned}
$$

The forward part for the coding theorem of $SMDCS[3,4,(1,3,4)]$ and $SMDCS$ $[3,5,(1,3,5)]$ are prove by showing the equivalence of $\mathbf{r_{35135}}$ and $\mathbf{R_{35135}}$ and the equivalence of $\mathbf{r_{35135}}$ and $\mathbf{R_{35135}}$ and are omitted. The converse part for $SMDCS$ $[3,4,(1,3,4)]$ are relatively easy and are also omitted. Here we prove the converse part of $SMDCS[3,5,(1,3,5)]$ by means of $\mathbf{R_{35135}}$. For any admissible 5-tuple $(R_1, R_2, R_3, R_4, R_5)$, obviously (3.23) are satified. (3.24) and (3.25) are obtained by invoking Theorem (1) and putting $X^1 = X, X^3 = Y$ and $X^5 = Z$ and degenerate the rest data streams. We just prove one case for each of (3.26) and (3.27) since the others are symmetrical and are satisfied similarly.

$$\begin{aligned}
& 2R_1 + R_2 + R_3 + R_4 + R_5 \\
\geq\ & n^{-1}(2H(S_1) + H(S_2) + H(S_3) + H(S_4) + H(S_5)) \\
=\ & n^{-1}(2H(S_1, X_1^n) + H(S_2, X_1^n) + H(S_3, X_1^n) + H(S_4, X_1^n) + H(S_5, X_1^n)) \\
\geq\ & n^{-1}(H(S_1, S_2, S_3, Y_1^n | X_1^n) + H(S_1, S_4, S_5, Y_1^n | X_1^n)) + 6H(X) \\
=\ & n^{-1}(H(S_1, S_2, S_3, S_4, S_5, Z_1^n | X_1^n, Y_1^n)) + 6H(X) + 2H(Y) \\
\geq\ & n^{-1}(H(S_1, S_2, S_3, S_4, S_5 | X_1^n, Y_1^n, Z_1^n)) + 6H(X) + 2H(Y) + H(Z) \\
\geq\ & 6H(X) + 2H(Y) + H(Z)
\end{aligned}$$

$$\begin{aligned}
& 3R_i + 3R_j + R_k + R_l + R_m \\
\geq\ & n^{-1}(3H(S_1) + 3H(S_2) + H(S_3) + H(S_4) + H(S_5)) \\
=\ & n^{-1}(3H(S_1, X_1^n) + 3H(S_2, X_1^n) + H(S_3, X_1^n) + H(S_4, X_1^n) + H(S_5, X_1^n)) \\
\geq\ & n^{-1}(H(S_1, S_2, S_3, Y_1^n | X_1^n) + H(S_1, S_2, S_4, Y_1^n | X_1^n)) + H(S_1, S_2, S_5, Y_1^n | X_1^n)) + 9H(X) \\
\geq\ & n^{-1}(H(S_1, S_2, S_3, S_4, S_5, Z_1^n | X_1^n, Y_1^n)) + 6H(X) + 3H(Y) \\
\geq\ & 9H(X) + 3H(Y) + H(Z)
\end{aligned}$$

The extreme points for $\mathbf{R_{34134}}$ and $\mathbf{R_{35135}}$ are listed below.

SMDCS	extreme points of coding rate region
$[3,4,(1,3,4)]$	$(H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3} + H(Z))$ $(\times 4)$
	$(H(X), H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2} + H(Z))$ $(\times 12)$
	$(H(X), H(X), H(X) + H(Y), H(X) + H(Y) + H(Z))$ $(\times 12)$
$[3,5,(1,3,5)]$	$(H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3}, H(X) + \frac{H(Y)}{3} + H(Z))$ $(\times 5)$
	$(H(X), H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2}, H(X) + \frac{H(Y)}{2} + H(Z))$ $(\times 20)$
	$(H(X), H(X), H(X) + H(Y), H(X) + H(Y), H(X) + H(Y) + H(Z))$ $(\times 20)$

3.5 SMDCS[4,4,(1,2,3,4)]

Now we come to the most complicated case of SMDCS in the thesis, $SMDCS[4,4,(1,2,3,4)]$ the configuration of which is shown in Figure 3.6. Assume the 4 independent data streams are $\{W_k\}, \{X_k\}, \{Y_k\}, \{Z_k\}$.

Also, every two encoders in $SMDCS[4,4,(1,2,3,4)]$ and the decoders which include only the two encoders as their fan forms a $SMDCS[2,2,(1,2)]$. Every three encoders in $SMDCS[4,4,(1,2,3,4)]$ and the decoders which include only the three encoders as their fan forms a $SMDCS[3,3,(1,2,3)]$.

The coding rate region described by subrate constraints, denoted as $\mathbf{r}_{441234}$ is $\{(R_1, R_2, R_3, R_4):$

$$R_i = r_i^w + r_i^x + r_i^y + r_i^z \quad \text{for } 1 \le i \le 4 \tag{3.29}$$

and

$$r_i^w \ge H(W) \quad \text{for } 1 \le i \le 4 \tag{3.30}$$

$$r_i^x + r_j^x \ge H(X) \quad \text{for } 1 \le i < j \le 4 \tag{3.31}$$

$$r_i^y + r_j^y + r_k^y \ge H(Y) \quad \text{for } 1 \le i < j < k \le 4 \tag{3.32}$$

$$r_1^z + r_2^z + r_3^z + r_4^z \ge H(Z) \quad \} \tag{3.33}$$

Define $\pi(1234)$ as the set of all permutation of the sequence (1,2,3,4). The coding rate region described by rate constraints, denoted as $\mathbf{R}_{441234}$, is given by $\{(R_1, R_2, R_3, R_4):$

$$R_i \ge H(W) \quad \text{for } 1 \le i \le 4 \tag{3.34}$$

$$R_i + R_j \ge 2H(W) + H(X) \quad \text{for } 1 \le i < j \le 4 \tag{3.35}$$

$$R_i + R_j + R_k \ge 3H(W) + \frac{3}{2}H(X) + H(Y) \tag{3.36}$$

$$2R_i + R_j + R_k \ge 4H(W) + 2H(X) + H(Y) \tag{3.37}$$

$$R_i + 2R_j + R_k \ge 4H(W) + 2H(X) + H(Y) \tag{3.38}$$

$$R_i + R_j + 2R_k \ge 4H(W) + 2H(X) + H(Y) \tag{3.39}$$

$$R_1 + R_2 + R_3 + R_4 \ge 4H(W) + 2H(X) + \frac{4}{3}H(Y) + H(Z) \tag{3.40}$$

$$2R_i + 2R_j + R_k + R_l \ge 6H(W) + 3H(X) + 2H(Y) + H(Z) \tag{3.41}$$

$$3R_i + R_j + R_k + R_l \ge 6H(W) + 3H(X) + \frac{3}{2} + H(Z) \tag{3.42}$$

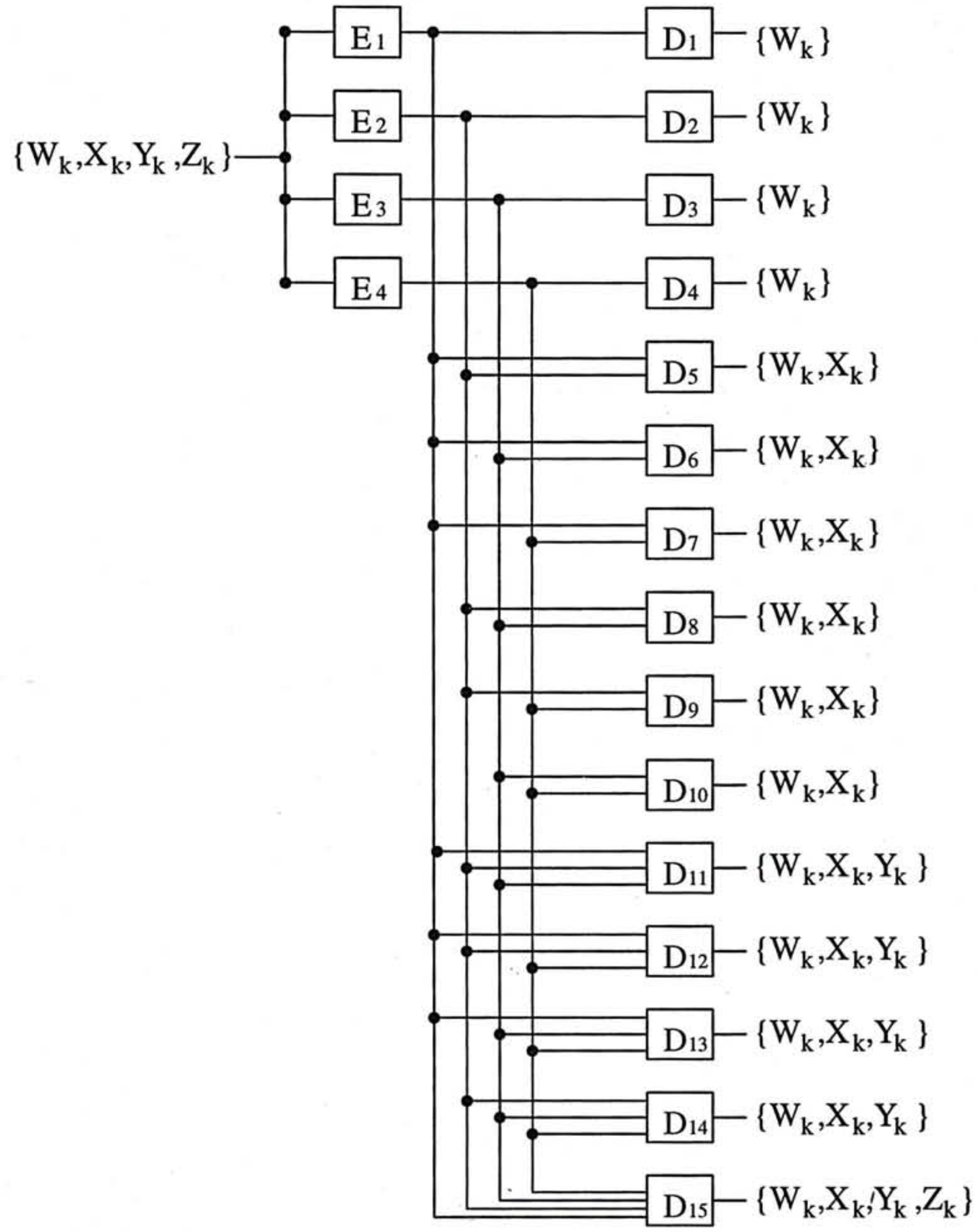


Figure 3.6: Configuration of SMDCS[4,4,(1,2,3,4)]

$$
\begin{aligned}
3R_i + 2R_j + 2R_k + 2R_l &\geq 7H(W) + \frac{9}{2}H(X) + 3H(Y) + 2H(Z) && (3.43)\\
4R_i + 2R_j + R_k + R_l &\geq 8H(W) + 4H(X) + 2H(Y) + H(Z) && (3.44)\\
&\quad \forall\ (i,j,k,l) \in \pi(1,2,3,4) \qquad \}
\end{aligned}
$$

Note that for (3.34) there are 4 constraints, for (3.35) there are 6 constraints, for (3.36) there are 4 constraints, for (3.37) , (3.38) and (3.39) there are 4 constraints repectively, for (3.41) there are 6 constraints, for (3.42) there are 4 constraints, for (3.43) there are 4 constraints, for (3.44) there are 12 constraints. Including (3.40), there are totally 53 rate constraints describing the set.

Theorem 4 *The admissible coding rate region of $SMDCS[4,4,(1,2,3,4)]$ is given by $\mathbf{R_{441234}}$, and superposition is the optimal coding scheme.*

Proof

The equivalence of the two sets $\mathbf{r_{441234}}$ and $\mathbf{R_{441234}}$ is proved in the Appendix A. $\mathbf{R_{441234}}$ is a subset of admissible coding rate region for $SMDCS[4,4,(1,2,3,4)]$ since superposition is a feasible coding scheme. Conversely, for any admissible 4-tuple (R_1, R_2, R_3, R_4), it is easily seen that (3.34) is satisfied for $1 \leq i \leq 4$. Since any two encoders forms a $SMDCS[2,2,(1,2)]$, any three encoders forms a $SMDCS[3,3,(1,2,3)]$. Moreover $SMDCS[2,2,(1,2)]$, $SMDCS[3,3,(1,2,3)]$ and $SMDCS[4,4,(1,2,3,4)]$ are all special cases of $SMDCS\ [m,m,(1,2,...,m)]$, so (3.35), (3.36) and (3.40) are satisfied by invoking Theorem (1). We will just prove one case in each of (3.41), (3.42),(3.43) and (3.44). The rest are proved similarily since they have symmetrical structures. In the following, (W_1^n) are functions of S_i for $1 \leq i \leq 4$, (X_1^n) are functions of (S_i, S_j) for $1 \leq i < j \leq 4$, (Y_1^n) are functions of (S_i, S_j, S_k) for $1 \leq i < j < k \leq 4$, (Z_1^n) are functions of (S_1, S_2, S_3, S_4).

$$
\begin{aligned}
& 2R_1 + 2R_2 + R_3 + R_4 \\
\geq\ & n^{-1}(2H(S_1) + 2H(S_2) + H(S_3) + H(S_4)) \\
=\ & n^{-1}(2H(S_1, W_1^n) + 2H(S_2, W_1^n) + H(S_3, W_1^n) + H(S_4, W_1^n)) \\
=\ & n^{-1}(2H(S_1|W_1^n) + 2H(S_2|W_1^n) + H(S_3|W_1^n) + H(S_4|W_1^n)) + 6H(W) \\
\geq\ & n^{-1}(H(S_1, S_2, X_1^n|W_1^n) + H(S_1, S_3, X_1^n|W_1^n) + H(S_2, S_4, X_1^n|W_1^n)) + 4H(W)
\end{aligned}
$$

$$
\begin{aligned}
&= n^{-1}(H(S_1,S_2,|W_1^n,X_1^n)+H(S_1,S_3|W_1^n,X_1^n)+H(S_2|W_1^n,X_1^n)\\
&\quad +H(S_4|S_2,W_1^n,X_1^n))+6H(W)+3H(X)\\
&\geq n^{-1}(S_1,S_2,S_4,Y_1^n|W_1^n,X_1^n)+H(S_1,S_2,S_3,Y_1^n|W_1^n,X_1^n))+6H(W)+3H(X)\\
&= n^{-1}(S_1,S_2,S_4|W_1^n,X_1^n,Y_1^n)+H(S_1,S_2,S_3|W_1^n,X_1^n,Y_1^n))\\
&\quad +6H(W)+3H(X)+2H(Y)\\
&\geq n^{-1}(H(S_1,S_2,S_3,S_4,Z_1^n|W_1^n,X_1^n,Y_1^n))+6H(W)+3H(X)+2H(Y)\\
&\geq 6H(W)+3H(X)+2H(Y)+H(Z)
\end{aligned}
$$

Also,

$$
\begin{aligned}
&3R_1+R_2+R_3+R_4\\
&\geq n^{-1}(3H(S_1)+H(S_2)+H(S_3)+H(S_4))\\
&= n^{-1}(3H(S_1,W_1^n)+H(S_2,W_1^n)+H(S_3,W_1^n)+H(S_4,W_1^n))\\
&= n^{-1}(3H(S_1|W_1^n)+H(S_2|W_1^n)+H(S_3|W_1^n)+H(S_4|W_1^n))+6H(W)\\
&\geq n^{-1}(H(S_1,S_2,X_1^n|W_1^n)+H(S_1,S_3,X_1^n|W_1^n)+H(S_1,S_4,X_1^n|W_1^n))+6H(W)\\
&= n^{-1}(H(S_1,S_2|W_1^n,X_1^n)+H(S_1,S_3|W_1^n,X_1^n)+H(S_1,S_4|W_1^n,X_1^n))\ +6H(W)+3H(X)\\
&\geq n^{-1}(\frac{1}{2}H(S_1,S_2,S_3,Y_1^n|W_1^n,X_1^n)+\frac{1}{2}H(S_1,S_2,S_4,Y_1^n|W_1^n,X_1^n)\\
&\quad +\frac{1}{2}H(S_1,S_3,S_4,Y_1^n|W_1^n,X_1^n))+6H(W)+3H(X)\\
&= n^{-1}(\frac{1}{2}H(S_1,S_2,S_3|W_1^n,X_1^n,Y_1^n)+\frac{1}{2}H(S_1,S_2,S_4|W_1^n,X_1^n,Y_1^n)\\
&\quad +\frac{1}{2}H(S_1,S_3,S_4|W_1^n,X_1^n,Y_1^n))+6H(W)+3H(X)+\frac{3}{2}H(Y)\\
&= n^{-1}(\frac{1}{2}H(S_1,S_2,S_3|W_1^n,X_1^n,Y_1^n)+\frac{1}{2}H(S_1,S_2,S_4|W_1^n,X_1^n,Y_1^n)\\
&\quad +\frac{1}{2}H(S_4|W_1^n,X_1^n,Y_1^n)+\frac{1}{2}H(S_1,S_3|W_1^n,X_1^n,Y_1^n,S_4))+6H(W)+3H(X)+\frac{3}{2}H(Y)\\
&\geq n^{-1}(H(S_1,S_2,S_3,S_4,Z_1^n|W_1^n,X_1^n,Y_1^n))+6H(W)+3H(X)+\frac{3}{2}H(Y)\\
&\geq 6H(W)+3H(X)+\frac{3}{2}H(Y)+H(Z)
\end{aligned}
$$

$$3R_1+2R_2+2R_3+2R_4$$

$$
\begin{aligned}
&\geq n^{-1}(3H(S_1)+2H(S_2)+2H(S_3)+2H(S_4)) \\
&= n^{-1}(3H(S_1,W_1^n)+2H(S_2,W_1^n)+2H(S_3,W_1^n)+2H(S_4,W_1^n)) \\
&= n^{-1}(3H(S_1|W_1^n)+2H(S_2|W_1^n)+2H(S_3|W_1^n)+2H(S_4|W_1^n))+9H(W) \\
&\geq n^{-1}(H(S_1,S_2,X_1^n|W_1^n)+H(S_1,S_3,X_1^n|W_1^n)+H(S_1,S_4,X_1^n|W_1^n) \\
&\quad +\frac{1}{2}H(S_2,S_3,X_1^n|W_1^n)+\frac{1}{2}H(S_2,S_4,X_1^n|W_1^n)+\frac{1}{2}H(S_3,S_4,X_1^n|W_1^n))+9H(W) \\
&= n^{-1}[(\frac{1}{2}(H(S_1,S_2|W_1^n,X_1^n)+H(S_1,S_3|W_1^n,X_1^n)+H(S_2,S_3|W_1^n,X_1^n)) \\
&\quad +\frac{1}{2}(H(S_1,S_3|W_1^n,X_1^n)+H(S_1,S_4|W_1^n,X_1^n)+H(S_3,S_4|W_1^n,X_1^n)) \\
&\quad +\frac{1}{2}(H(S_1,S_2|W_1^n,X_1^n)+H(S_1,S_4|W_1^n,X_1^n)+H(S_2,S_4|W_1^n,X_1^n))]+\frac{9}{2}H(X)+9H(W) \\
&\geq n^{-1}(H(S_1,S_2,S_3,Y_1^n|W_1^n,X_1^n)+H(S_1,S_3,S_4,Y_1^n|W_1^n,X_1^n) \\
&\quad +H(S_1,S_2,S_4,Y_1^n|W_1^n,X_1^n))+\frac{9}{2}H(X)+9H(W) \\
&= n^{-1}(H(S_1,S_2,S_3|W_1^n,X_1^n,Y_1^n)+H(S_4|S_1,S_3,W_1^n,X_1^n,Y_1^n) \\
&\quad +H(S_1,S_3|W_1^n,X_1^n,Y_1^n)+H(S_1,S_2,S_4,Y_1^n|W_1^n,X_1^n))+\frac{9}{2}H(X)+3H(Y) \\
&\geq n^{-1}2H(S_1,S_2,S_3,S_4,Z_1^n|W_1^n,X_1^n,Y_1^n))+3H(Y)+\frac{9}{2}H(X)+9H(W) \\
&\geq 9H(W)+\frac{9}{2}H(X)+3H(Y)+2H(Z)
\end{aligned}
$$

$$
\begin{aligned}
&4R_1+2R_2+R_3+R_4 \\
&\geq n^{-1}(4H(S_1)+2H(S_2)+H(S_3)+H(S_4)) \\
&= n^{-1}(4H(S_1,W_1^n)+2H(S_2,W_1^n)+H(S_3,W_1^n)+H(S_4,W_1^n)) \\
&= n^{-1}(4H(S_1|W_1^n)+2H(S_2|W_1^n)+H(S_3|W_1^n)+H(S_4|W_1^n))+8H(W) \\
&\geq n^{-1}(2H(S_1,S_2,X_1^n|W_1^n)+H(S_1,S_3,X_1^n|W_1^n)+H(S_1,S_4,X_1^n|W_1^n))+8H(W) \\
&\geq n^{-1}(H(S_1,S_2,S_3,Y_1^n|W_1^n,X_1^n)+H(S_1,S_2,S_4,Y_1^n|W_1^n,X_1^n))+8H(W)+4H(X) \\
&\geq n^{-1}(H(S_1,S_2,S_3,S_4,Z_1^n|W_1^n,X_1^n,Y_1^n))+8H(W)+4H(X)+2H(Y) \\
&\geq 8H(W)+4H(X)+2H(Y)+H(Z)
\end{aligned}
$$

So superposition is the optimal coding scheme. □

By degnerating streams $\{W_k\}, \{X_k\}$ or $\{Y_k\}$, we can specialize $SMDCS[4,4,(1,2,3,4)]$

to 3 special cases which are $SMDCS[3,4,(2,3,4)]$, $SMDCS[3,4,(1,2,4)]$ or $SMDCS$ $[3,4,(1,3,4)]$ respectively and the coding rate regions of the 2 latter cases are special cases of $SMDCS[3,m,(1,2,m)]$, $SMDCS[3,m,(1,3,m)]$, the coding rate regions of which are stated in the previous sections. Consistency with the previous results is obvious. As one may expect, the set of constraints describing $\mathbf{R_{34124}}, \mathbf{R_{34134}}$ and $\mathbf{R_{34123}}$ are the same as some constraints in that of R_{441234}. By looking at the coding rate regions of these three special cases, we can make more sense out of the set of constraints describing the $\mathbf{R_{441234}}$ which can be viewed as a *superset of all the rate constraints of its special cases.*

Note that by degenerating stream $\{Z_k\}$ we reduce $SMDCS[4,4,(1,2,3,4)]$ to $SMDCS[3,4,(1,2,3)]$ which is a special case of $SMDCS[3,m,(1,2,3)]$. In this case (3.40), (3.41), (3.42), (3.43) and (3.44) of $\mathbf{R_{441234}}$ are implied by the rest of the constraints so are redundant and are deleted. As a result we obtained the the admissible codng rate region for $SMDCS[3,4,(1,2,3)]$, $\mathbf{R_{34123}}$. The result is consistent with the conjecture in Section 3.3. Also by degerating $\{X_k\}$ and $\{Y_k\}$ altogether we obtain $SMDCS[2,4,(1,2)]$, by degerating $\{X_k\}$ and $\{Z_k\}$ altogether we obtain $SMDCS[2,4,(1,3)]$ and by degerating $\{Y_k\}$ and $\{Z_k\}$ altogether we obtain $SMDCS[2,4,(1,2)]$. They are special cases of $SMDCS[2,m,(1,k)]$. The results here are consistent with the conjecture in Section 3.2.

Chapter 4

Convex Analysis of Coding Rate Region of DCS

4.1 Introduction

In this chapter we are going to analyze the coding rate regions of MDCS's induced by superposition. Two representations of coding rate regions, $\mathbf{r}_{\mathbf{sp}}$ and $\mathbf{R}_{\mathbf{sp}}$ are studied throughout the thesis. Two operations involving them are important in obtaining the results in this thesis. On one hand we are to explicitly construct the rate constraints from the given set of subrate constraints. On the other hand a set given in terms of rate constraints, we are to find simple method to prove that it is equivalent to a set given in terms of subrate constraints. The first operation is important as we want to prove the optimality of superposition, the converse part of the MDCS coding must be proved to be true. We are unable to do so without a complete knowledge of the $\mathbf{R}_{\mathbf{sp}}$ in terms of rate constraints. Coding rate region are in fact positive polyhedral sets with special properties. By finding the convex hull of the set formed by summing the extreme points of the subrate constraints, we are able to find out some of the rate constraints which are difficult to discover otherwise. This will be explained in this chapter. For the second operation, as it is nontrivial to construct a set $\mathbf{R}_{\mathbf{sp}}$ described by a set of rate constraints from a set $\mathbf{r}_{\mathbf{sp}}$ of subrate constraints which are obvious once the MDCS is defined and the construction process are difficult to trace, we want to prove the equivalence of the two by simple method. We can provide a primitive algorithm to explicitly check whether any two given polyhedral sets $\mathbf{r}_{\mathbf{sp}}$ and $\mathbf{R}_{\mathbf{sp}}$ are equivalent. It is done by enumerating all the extreme points of the rate constraints and checking whether they are the sum of the extreme points of the correponding subrate constraints.

Also it is interesting to ask whether an admissible rate tuple of an MDCS is achievable by applying superposition. This involves checking whether we can decompose the tuple into subrate tuples satisfying the subrate constraints defined by the MDCS and if it is so we would like to know how we can actually allocate the subrate in each component of the tuple, that is to allocate coding rate of each encoder to encode one of the independent streams. We also provide an simple algorithm to do so.

The theoretical foundations for the above issues are explained in detail in this chapter and we invoke mathematical tools from convex analysis of polyhedral sets, computational geometry and linear programming to solve the problems.

4.2 Polyhedral Sets

To facilitate further discussion, we review some concepts in convex set analysis. In the following a point in $\mathbb{R}^n$ is represented by a column n-vector.

1) A set $\mathbf{X}$ in $\mathbb{R}^n$ is called *convex set* if given any two points $\mathbf{x_1}$ and $\mathbf{x_2}$ in $\mathbf{X}$ then $\mathbf{x} = \lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}$ is also in $\mathbf{X}$ for all λ in $[0, 1]$. In the following discussion all set denoted by $\mathbf{X}$ is convex unless specified otherwise.

2) A point $\mathbf{x}$ in $\mathbf{X}$ is a *convex combination* of points of $\mathbf{X}$ if there exists a finite set of points $\{\mathbf{x_i}, i = 1, 2, \ldots, k\}$ in $\mathbf{X}$ such that $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x_i}$ and $\sum_{i=1}^{k} \lambda_i = 1$. The *convex hull* of $\mathbf{X}$, denoted by $conv(\mathbf{X})$, is the set of all points that are convex combination of points in $\mathbf{X}$. $conv(\mathbf{X}) = \mathbf{X}$ if and only if $\mathbf{X}$ is convex.

3) A point $\mathbf{x}$ in $\mathbf{X}$ is called an *extreme point* of $\mathbf{X}$ if it cannot be represented as a convex combination of 2 distinct points in $\mathbf{X}$. In other words, if $\mathbf{x} = \lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}$ with $\lambda \in (0, 1)$ and $\mathbf{x_1}, \mathbf{x_2} \in \mathbf{X}$, then $\mathbf{x} = \mathbf{x_1} = \mathbf{x_2}$. A non-zero vector $\mathbf{d} \in \mathbf{X}$ is called a *direction* of $\mathbf{X}$ if for each $\mathbf{x} \in \mathbf{X}$, the point $\mathbf{x} + \mu\mathbf{d}$ is also in $\mathbf{X}$ $\forall \mu \geq 0$. An *extreme direction* of $\mathbf{X}$ is a direction that cannot be represented by the convex combination of two distinct direction. For a unique representation, the set of extreme directions $\mathbf{D}$ are normalized such that $\sum_{i=1}^{n} d_i = 1$ for each $\mathbf{d} \in \mathbf{D}$ where d_i is the ith component of $\mathbf{d}$. Any direction of the set $\mathbf{X}$ can be represented as a *linear combination* of the set of extreme directions of $\mathbf{X}$.

4) A hyperplane in $\mathbb{R}^n$ is a set of form $\{\mathbf{x} : \mathbf{a}^\mathbf{T}\mathbf{x} = \mathbf{k}\}$, where $\mathbf{a}$ is a nonzero column n-vector in $\mathbb{R}^n$. A hyperplane divide the $\mathbb{R}^n$ into two regions, each of them is called a *halfspace*. A halfspace $H \in \mathbb{R}^n$ is represented by a linear inequality $\{\mathbf{x} : \mathbf{a}^\mathbf{T}\mathbf{x} \geq k\}$. A polyhedral set $\mathbf{X}$ is a convex set formed by the intersection of a finite number of halfspace represented by the system of inequalities $\{\mathbf{a_i^T}\mathbf{x} \geq b_i, i = 1, \ldots, m\}$. Equivalently $\mathbf{X}$ is represented by $\{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ where $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b}$ as a m row vector. The hyperplane corresponding to one of the inequalities in the system is called the *defining hyperplane* of the set. We insist on the linear independence of the set of vectors $[\mathbf{a_i^T}, b_i]$, $i = 1, 2, \ldots, m$, i.e. none is a linear combination of the others in the set for a minimum representation. A bounded polyhedral set $\mathbf{P}$ is called a *polytope* and each element $\mathbf{x} \in \mathbf{P}$ satisfy $|\mathbf{x}| < K$ for a finite number K. In the following a polyhedral set is unbounded unless specified otherwise. It is easily seen that a polyhedral set is a convex set.

5) An alternative representation of polyhedral set is given by the Representation Theorem [2] [4].

Theorem 5 *Any element in a polyhedral set can be repersented by the convex combination of the extreme points of the set and the non-negative linear combination of its extreme directions. i.e., let $\mathbf{X}$ be a nonempty polyheral set with extreme points $\mathbf{x_i}, i = 1, \ldots, k$ and extreme directions $\mathbf{d_j}, j = 1, \ldots, l$, then*

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x_i} + \sum_{j=1}^{l} \mu_j \mathbf{d_j} \quad \forall \mathbf{x} \in \mathbf{X} \tag{4.1}$$

and

$$\sum_{i=1}^{k} \lambda_i = 1 \;\; \lambda_i \geq 0 \;\; i = 1, \ldots, k$$

$$\mu_i \geq 0 \;\; j = 1, \ldots, l$$

It is easy to see that a polytope $\mathbf{P}$ has no dircetions and the representation of $\mathbf{P}$ is reduced to

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x_i} \;\; \forall\, \mathbf{x} \in \mathbf{P} \text{ and } \sum_{i=1}^{k} \lambda_i = 1$$

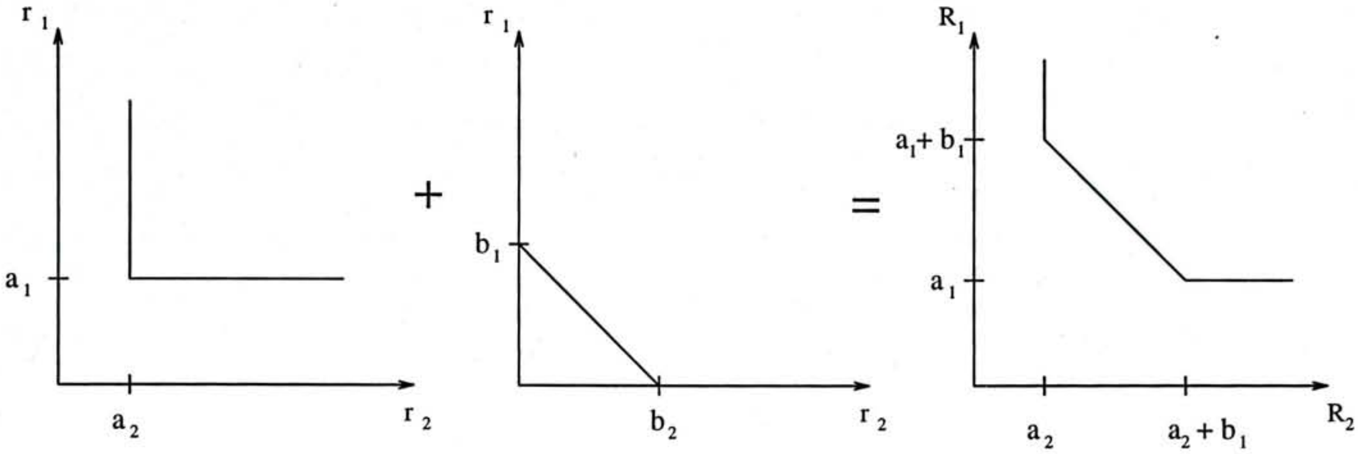


Figure 4.1: $\mathbf{r_1} + \mathbf{r_2} = \mathbf{R}$

from the representation in (4.1), we can think of a polyhedral set as a set formed by adding a linear combination of a fixed set of points (extreme directions in this case) to the points in a convex polytope. For any polyhedral set with a set of extreme points and extreme directions there is a *unique* polytope determined by the same set of extreme points.

4.3 Addition of Polyhedral Sets

For two sets $\mathbf{X}$ and $\mathbf{Y}$ in $\mathbb{R}^n$, define the addition of $\mathbf{X}$ and $\mathbf{Y}$ as $\mathbf{X} + \mathbf{Y} = \mathbf{Z}$ where the *sum*, $\mathbf{Z}$ is the set $\{\mathbf{z} = \mathbf{x} + \mathbf{y} : \mathbf{x} \in \mathbf{X}, \mathbf{y} \in \mathbf{Y}\}$. $\mathbf{X}$ and $\mathbf{Y}$ are called *summands* of $\mathbf{Z}$. Addition are defined for sets of the same dimension. The following properties hold in set addition,

$$\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$$

$$(\mathbf{W} + \mathbf{X}) + \mathbf{Y} = \mathbf{W} + (\mathbf{X} + \mathbf{Y})$$

For $\mathbf{W}, \mathbf{X}, \mathbf{Y} \in \mathbb{R}^n$

Example 4.1

Define $\mathbf{r_1} \in \mathbb{R}^2$ as $\{(r_1, r_2) : \ r_1 \geq a_1,\ r_2 \geq a_2\}$ and $\mathbf{r_2} \in \mathbb{R}^2$ as $\{(r_1, r_2) : \ r_i \geq 0, \text{ for } \ i = 1, 2.\ r_1 + r_2 \geq b_1\}$. $\mathbf{R} = \mathbf{r_1} + \mathbf{r_2}$ is represented by $\{(R_1, R_2) : \ R_1 \geq a_1,\ R_2 \geq a_2,\ R_1 + R_2 \geq a_1 + a_2 + b_1\}$. Note that the extreme points of $\mathbf{R}$ can be represented as sum of extreme points of $\mathbf{r_1}$ and $\mathbf{r_2}$. The three sets are shown in Figure 4.1.

We find that the representation in terms of inequalities has little operational meaning as far as the addition of set is concerned. So we turn to the repersentation in terms of extreme points and extreme directions for more insight. We consider the addition of two polyheral sets $\mathbf{X}$ and $\mathbf{Y}$ with different sets of extreme points $\mathbf{E_x} = \{\mathbf{x_i}, i = 1, 2, \ldots l\}$ and $\mathbf{E_y} = \{\mathbf{y_j}, j = 1, 2, \ldots, m\}$ but with the same set of extreme direction $\mathbf{D} = \{\mathbf{d_k}, k = 1, 2, \ldots, n\}$. Let $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$. By the Representation Theorem. $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ can be represented as

$$\begin{aligned}
\mathbf{x} &= \sum_{i=1}^{l} \alpha_i \mathbf{x_i} + \sum_{k=1}^{n} \lambda_k \mathbf{d_k} \\
\mathbf{y} &= \sum_{j=1}^{m} \beta_j \mathbf{y_j} + \sum_{k=1}^{n} \mu_k \mathbf{d_k} \\
\mathbf{z} &= \sum_{i=1}^{l} \sum_{j=1}^{m} \alpha_i \beta_j (\mathbf{x_i} + \mathbf{y_j}) + \sum_{k=1}^{n} (\lambda_k + \mu_k) \mathbf{d_k} && (4.2) \\
&= \sum_{q=1}^{p} \gamma_q \mathbf{z_q} + \sum_{k=1}^{n} \nu_k \mathbf{d_k} && (4.3) \\
& \forall\ \mathbf{x} \in \mathbf{X}\ \ \mathbf{y} \in \mathbf{Y}\ \ \mathbf{z} \in \mathbf{Z}, && (4.4) \\
& \sum_{i=1}^{l} \alpha_i = \sum_{j=1}^{m} \beta_j = \sum_{q=1}^{p} \gamma_q = 1, && (4.5) \\
& \alpha_i, \beta_j, \lambda_k, \mu_k, \gamma_q, \nu_k \geq 0. && (4.6)
\end{aligned}$$

$\mathbf{E_z} = \{\mathbf{z_q}, q = 1, 2, \ldots, p\}$ is the set of extreme points of $\mathbf{Z}$. From (4.3) it is quite obvious that the sum of polyhedral sets is a polyhedral set. Also if we delete all the terms involving directions we also have the sum of polytopes is a polytope. It is quite intuitive that the set $\mathbf{E_z}$ is a subset of the set $\mathbf{E_z} = \{\mathbf{z} : \mathbf{z} = \mathbf{x_i} + \mathbf{y_j}, i = 1, 2, \ldots l, j = 1, 2, \ldots, m\}$. We state this explicitly in the following theorem.

Theorem 6 *Every extreme point* $\mathbf{z}$ *of the convex set* $\mathbf{Z}$ *where* $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ *has a unique representation of* $\mathbf{z} = \mathbf{x} + \mathbf{y}$ *where* $\mathbf{x}$ *is a extreme point of* $\mathbf{X}$ *and* $\mathbf{y}$ *is an extreme point of* $\mathbf{Y}$.

proof

First we prove that representation of $\mathbf{z}$ is unique. Suppose $\mathbf{z}$ is represented as $\mathbf{x_1} + \mathbf{y_1}$ and $\mathbf{x_2} + \mathbf{y_2}$ where $\mathbf{x_1}, \mathbf{x_2} \in \mathbf{X}$ and $\mathbf{y_1}, \mathbf{y_2} \in \mathbf{Y}$, then $\mathbf{x_1} + \mathbf{y_2} \in \mathbf{Z}$ and

$\mathbf{x_2} + \mathbf{y_1} \in \mathbf{Z}$. Then

$$\mathbf{z} = \frac{1}{2}(\mathbf{x_1} + \mathbf{y_1}) + \frac{1}{2}(\mathbf{x_2} + \mathbf{y_2}) = \frac{1}{2}(\mathbf{x_1} + \mathbf{y_2}) + \frac{1}{2}(\mathbf{x_2} + \mathbf{y_1})$$

which contradict the extremeness of $\mathbf{z}$ unless $\mathbf{x_1} + \mathbf{y_2}$ and $\mathbf{x_2} + \mathbf{y_1}$ are the same point. But in that case $\mathbf{x_1} + \mathbf{y_2} = \mathbf{x_2} + \mathbf{y_1}$ together with $\mathbf{x_1} + \mathbf{y_1} = \mathbf{x_2} + \mathbf{y_2}$, we have $\mathbf{x_1} = \mathbf{x_2}$ and $\mathbf{y_1} = \mathbf{y_2}$. Thus the representation of $\mathbf{z}$ is unique.

Next suppose $\mathbf{x_1}$ is not an extreme point of $\mathbf{X}$ which is then represented as $\lambda \mathbf{a_1} + (1 - \lambda)\mathbf{a_2}$ where $\mathbf{a_1}, \mathbf{a_2} \in \mathbf{X}$, $\mathbf{a_1} \neq \mathbf{a_2}$ and $\lambda > 0$. Then

$$\mathbf{z} = \lambda \mathbf{a_1} + (1 - \lambda)\mathbf{a_2} + \mathbf{y_1} = \lambda(\mathbf{a_1} + \mathbf{y_1}) + (1 - \lambda)(\mathbf{a_2} + \mathbf{y_1})$$

Since $\mathbf{a_1} + \mathbf{y_1} \neq \mathbf{a_2} + \mathbf{y_1}$, this contradicts the extremeness of $\mathbf{z}$. So $\mathbf{x_1}$ must be an extreme point of $\mathbf{X}$. The same argument applies to $\mathbf{y_1}$. □

If we consider the case in which $\mathbf{Z} = \mathbf{X_1} + \mathbf{X_2} + \ldots + \mathbf{X_m}$ where $\mathbf{Z}, \mathbf{X_i} \in \mathbb{R}^n$ for $1 \leq i \leq m$ is convex and $\mathbf{X_i}$ has extreme point set $\mathbf{E_i}$. By adding the set one by one and applying Theorem 6 consecutively we have the following. Each extreme point $\mathbf{z_k}$ of the set $\mathbf{Z}$ is uniquely represented as,

$$\mathbf{z_k} = \mathbf{x_{1k}} + \mathbf{x_{2k}} + \ldots + \mathbf{x_{nk}} \quad \text{where } \mathbf{x_{ik}} \in \mathbf{E_i}$$

Now we focus on polytope, a bounded polyhedral sets. Since from the Representation Theorem all points in a polytope can be represented by convex combination of the extreme points of the polytope, it is obvious that given the set of extreme points of a polytope is a sufficient characterization of the polytope. Also we see that a polytope is also represented by a set of inequalities. Now the question is how we can relate one representation with the equivalent counterpart. We state without proof two facts in linear algebra which relates the 2 representations.

1. A extreme point of a convex polytope $\mathbf{X} \in \mathbb{R}^n$ is the intersection of n linearly independent defining hyperplanes of the set.

2. A defining hyperplane of a convex polytope $\mathbf{X} \in \mathbb{R}^n$ contains at least n extreme points.

When we consider the addition of two convex polytopes given by their corresponding set of ineqalities, we propose a method to find the sum which is another polytope.

1. Find the extreme points of each (summand) polytope.
2. Creat a set which is the sum of the extreme points of the two polytopes.
3. Find the set of extreme points of the sum from convex hull computation.
4. Convert the extreme point representation back to the inequalities representation.

How can we find tools to help us? In computational geometry [9] there are standard algorithms which can find the convex hull of a given set of points. The convex hull in this context is the polytopes that contain the set. The algorithms generally can

1. Enumerate all the extreme points in the set of input points.
2. identify which of the extreme points are contained in the same defining hyperplane of the set.

The equation of a hyperplane can be solved from the set of extreme points contained in the hyperplane, we can obtain the set of inequalities representing the polytope.

For the addition of the polyhedral set as in (4.2), before we actually find the sum which is another polyhedral set (which I don't know how to find all the inequalities in a few steps at the moment), first we reduce the problem to the addition of two coresponding polytopes uniquely determined by the two polyhedral sets and then find out the sum polytope (which I know how to find all the inequalities in a few steps now). We can discover a subset of the describing inequalities of the sum polyhedral set which is common to both the sum polytope and the sum polyhedral set. We will see from the coming examples that what we can find from the sum polytope actually cover most of the nontrivial inequalities in the sum polyhedral set. Also the rest inequalities describing the sum polyhedral set can be deduced from the summands.

Now a convex polyhedral set represented as (4.1) is an unbounded set since every point in the convex polytope added by a linear combination of the directions is also in the set. So some inequalities of the polytope are not valid constraints for the corresponding polyhedral given the set of directions. We are interested in finding which of the inequalities is common to both the polytope and the polyhedral set. Assume the polyhedral set has extreme direction set $\{\mathbf{d_j}, j = 1, 2, \ldots, n\}$. For a valid inequality $\mathbf{a_i^T x} \geq b_i$ of the polyhedral set (note that we insist on the $\geq$ sign) in (4.1), if $\mathbf{x}$ satisfies the inequalities then $\mathbf{x}+\mathbf{d}$ must also satisfy the inequalities, where $\mathbf{d}$ is a direction of the set $\mathbf{X}$. Then,

$$\begin{aligned} \mathbf{a_i}^T(\mathbf{x}+\mathbf{d}) &= \mathbf{a_i}^T\mathbf{x}+\mathbf{a_i}^T\mathbf{d} \\ &\geq b_i \\ \Rightarrow \qquad \mathbf{a_i}^T\mathbf{d} &\geq 0 \\ \Rightarrow \quad \mathbf{a_i}^T(\textstyle\sum_{j=1}^n \mu_j \mathbf{d_j}) &\geq 0 \text{ where } \mu_j > 0 \text{ for } 1 \leq j \leq n \\ \Rightarrow \quad \textstyle\sum_{j=1}^n \mu_j(\mathbf{a_i}^T\mathbf{d_j}) &\geq 0 \text{ for } 1 \leq j \leq n \end{aligned}$$

$$\Rightarrow \qquad \mathbf{a_i}^T\mathbf{d_j} \geq 0 \quad \text{for } 1 \leq j \leq n \tag{4.7}$$

The last inequalities must hold since $\mu_j, j = 1, 2, \ldots, n$ can be arbitrarily large. Since (4.7) must be satisfied for the set of the inequalities describing the polyhedral set, so by checking (4.7) for all $\mathbf{a_i}$ with the set of extreme directions we can eliminate all the invalid inequalities and remaining ones is the inequalities common to both the polyhedral set and the corresponding polytope.

In fact, for a polyhedral set represented by $\{\mathbf{x} : \ \mathbf{Ax} \geq \mathbf{b}\}$ the set of (normalized) directions are given by $\mathbf{D} = \{\mathbf{d} = (d_1, d_2, \ldots, d_n) : \mathbf{Ad} \geq 0, \sum_{i=1}^n d_i = 1\}$. Extreme directions are the extreme points of the set $\mathbf{D}$ [2]. In particular, for a *positive* polydedral set in $\mathbb{R}^n$ described by $\{\mathbf{x} : \ \mathbf{Ax} \geq \mathbf{b}\}$ where all the entries in $\mathbf{A}$ are also *positive*, $\mathbf{Ax} \geq 0$ can be reduced to $\mathbf{Ix} \geq 0$ where $\mathbf{I}$ is a $n \times n$ identity matrix. It is obvious that the set of extreme point of the set $\mathbf{D} = \{\mathbf{d} = (d_1, d_2, \ldots, d_n) : \mathbf{Id} \geq 0, \sum_{i=1}^n d_i = 1\}$ is $\{(1,0,0,\ldots,0),(0,1,0,0,\ldots,0),\ldots,(0,0,\ldots,1)\}$. All the coding rate regions in this paper are positive polyhedral sets and the coefficients of the inequalities are all positive. So the set of extreme directions are immediately determined once the *dimension* n of the set are given. When we check the set of inequalities with the set of extreme directions by (4.7), it turn out that in all the valid inequalities $\mathbf{a_i}^T\mathbf{x} \geq b_i$, $\mathbf{a_i}$ should be a *positive* vector.

Example 4.2

Consider the addition of 2 sets $\mathbf{r^x} + \mathbf{r^y} = \mathbf{R} \in \mathbb{R}^3$ where $\mathbf{r^x}$ is
$\{(r_1^x, r_2^x, r_2^x) : \ r_i^x \geq 0 \ \text{ for } 1 \leq i \leq 3, \ r_1^x + r_2^x \geq 2, \ r_2^x + r_3^x \geq 2 \ \}$
and $\mathbf{r^y}$ is $\{(r_1^y, r_2^y, r_3^y) : \ \ r_1^y + r_2^y + r_3^y \geq 3 \ \}$

The corresponding extreme point set $\mathbf{E_x}, \mathbf{E_y}$ and extreme direction set $\mathbf{D}$ is given by

$$\mathbf{E_x} \ = \ \{(0,2,0),(2,0,2)\}$$

$$\begin{aligned}
\mathbf{E_y} &= \{(3,0,0),(0,3,0),(0,0,3)\} \\
\mathbf{D} &= \{(1,0,0),(0,1,0),(0,0,1)\}
\end{aligned}$$

The set of extreme point sum $\mathbf{E_R}$ is given by

$$\{(3,2,0),\ (5,0,2),\ (0,5,0),\ (2,3,2),\ (0,2,3),\ (2,0,5)\}$$

By finding the convex hull of the sum of extreme points and express the halfspaces in terms of inequalities we have,

$$\begin{aligned}
R_1 + R_2 &\leq 5 \\
R_2 + R_3 &\leq 5 \\
R_1 + R_2 + R_3 &\leq 7 \\
R_1 + R_2 + R_3 &\geq 5 \\
R_1 + 2R_2 + R_3 &\geq 7
\end{aligned}$$

These are the inequalities describing the polytopes. Since what we obtained from the convex hull program is just hyperplanes equations, the inequality sign has to be obtained by putting a point which is not contained in the hyperplane from set $\mathbf{E_R}$ into the equation of the hyperplane. In fact the sum $\mathbf{R}$ of the 2 sets is $\{(R_1, R_2, R_3):$

$$\begin{aligned}
R_1 + R_2 &\geq 2 \\
R_2 + R_3 &\geq 2 \\
R_1 + R_2 + R_3 &\geq 5 \\
R_1 + 2R_2 + R_3 &\geq 7 \quad \}
\end{aligned}$$

The first two inequalities in the set is not given by the convex hull program which can just find the defining half space of the polytope containing the set of extreme point sum. They appear because of the addition of directions to the polytope creating new hyperplanes other than those of the original polytope. These two inequalities are readily obtained by inspection since it is the same as the inequalities in $\mathbf{r^x}$. It is obvious that the inequalities of the every summand must be a subset of inequalities describing the sum. This is a special result of the following theorem.

Theorem 7 *Suppose* $\mathbf{R} = \mathbf{r_1} + \mathbf{r_2} + \ldots + \mathbf{r_m}$ *and* $\mathbf{r_i} \in I\!R^{n+}$, *for* $1 \leq i \leq m$, *and* $\mathbf{r_i} = \{\mathbf{x} : \mathbf{A_i x} \geq \mathbf{b_i}\}$ *such that* $\mathbf{A_i} \geq \mathbf{0}$ *componentwise for* $1 \leq i \leq m$ *and* $\mathbf{R} =$

$\{\mathbf{x}: \ \mathbf{Ax} \geq \mathbf{b}\}$. *Also ,let*

$$\mathbf{r_{i_1}} + \mathbf{r_{i_2}} + \ldots + \mathbf{r_{i_k}} = \{\mathbf{x}: \ \mathbf{A_{i_1,i_2,\ldots,i_k}x} \geq \mathbf{b_{i_1,i_2,\ldots,i_k}}\} \text{ for } 1 \leq i_1 < i_2 < \ldots < i_k \leq m \tag{4.8}$$

Then the rows of $\mathbf{A_{i_1,i_2,\ldots,i_k}}$ *for* $1 \leq i_1 < i_2 < \ldots < i_k \leq m$ *is a subset of the rows of* $\mathbf{A}$. *(The row* $\mathbf{a}$ *is assumed to be equivalent to* $k\mathbf{a}$ *where* k *is constant, i.e. two inequalities are assumed to be identical if the relative ratio of the coefficients are equal.)*

Proof

First we note that for any two polyhedral sets $\mathbf{X_1}, \mathbf{X_2} \subset \mathbb{R}^{n+}$ and $\mathbf{X_1} = \{\mathbf{x}: \mathbf{B_1x} \geq \mathbf{c_1}\}$ and $\mathbf{X_2} = \{\mathbf{x}: \mathbf{B_2x} \geq \mathbf{0}\}$ where $\mathbf{B_1}, \mathbf{B_2} \geq \mathbf{0}$ componentwise, $\mathbf{X_1} + \mathbf{X_2} = \mathbf{X_1}$. Given we have found the representation of $\mathbf{R}$ as the sum of the sets $\mathbf{r_i}$ where $\mathbf{b_i}$ is greater than 0 for $1 \leq i \leq m$. Then we notice that the representation of $\mathbf{r_{i_1}} + \mathbf{r_{i_2}} + \ldots + \mathbf{r_{i_k}}$ is obtained by setting $\mathbf{b_i} = \mathbf{0}$ for $1 \leq i \leq m, i \notin \{i_1, i_2, \ldots, i_k\}$ in $\mathbf{Ax} \geq \mathbf{b}$ and $\mathbf{R} = \{\mathbf{x}: \ \mathbf{Ax} \geq \mathbf{b}\}$ is reduced to $\{\mathbf{x}: \ \mathbf{A_{i_1,i_2,\ldots,i_k}x} \geq \mathbf{b_{i_1,i_2,\ldots,i_k}}\}$ which is the result of deleting some of the redundant rows in $\mathbf{Ax} \geq \mathbf{b}$ after setting $\mathbf{b_i} = \mathbf{0}$ for $1 \leq i \leq m, i \notin \{i_1, i_2, \ldots, i_k\}$. This is a reduction process in which no new inequalities are created other than those already exist in $\mathbf{Ax} \geq \mathbf{b}$. Then the rows of $\mathbf{A_{i_1,i_2,\ldots,i_k}}$ must be some rows in $\mathbf{A}$ before setting $\mathbf{b_i} = \mathbf{0}$ for $1 \leq i \leq n, i \notin \{i_1, i_2, \ldots, i_k\}$. The arguments are true for $1 \leq k \leq m - 1$. □

Example 4.3

Consider the addition of 3 sets $\mathbf{r^x} + \mathbf{r^y} + \mathbf{r^z} = \mathbf{R} \in \mathbb{R}^3$ where

$\mathbf{r^x}$ is

$$\{(r_1^x, r_2^x, r_3^x): \ \ r_1^x \geq 1, r_2^x \geq 1, r_3^x \geq 1 \ \}$$

$\mathbf{r^y}$ is

$$\begin{aligned} \{(r_1^y, r_2^y, r_3^y): & \\ r_i^y &\geq 0 \text{ for } 1 \leq i \leq 3 \\ r_1^y + r_2^y &\geq 2 \\ r_1^y + r_3^y &\geq 2 \\ r_2^y + r_3^y &\geq 2 \quad \} \end{aligned}$$

and $\mathbf{r^z}$ is

$$\{(r_1^z, r_2^z, r_3^z): r_i^z \geq 0, \text{ for } 1 \leq i \leq 3, \ r_1^z + r_2^z + r_3^z \geq 3\}$$

The corresponding extreme point set $\mathbf{E_x}, \mathbf{E_y}, \mathbf{E_z}$ and the common extreme direction set $\mathbf{D}$ is given by

$$\begin{aligned}
\mathbf{E_x} &= \{(1,1,1)\} \\
\mathbf{E_y} &= \{(2,2,0),(2,0,2),(0,2,2)\} \\
\mathbf{E_z} &= \{(3,0,0),(0,3,0),(0,0,3)\} \\
\mathbf{D} &= \{(1,0,0),(0,1,0),(0,0,1)\}
\end{aligned}$$

The set of extreme points sum is given by

$$\{\ (6,3,1),(6,1,3),(4,3,3),(5,2,2),(3,6,1),(3,4,3)$$

$$(1,6,3),(2,5,2),(3,3,4),(3,1,6),(1,3,6),(2,2,5)\}$$

By finding the convex hull of the sum of extreme points and express the halfspaces in terms of inequalities we have,

$$\begin{aligned}
R_1 + R_2 &\geq 4 \\
R_1 + R_3 &\geq 4 \\
R_2 + R_3 &\geq 4 \\
2R_1 + R_2 + R_3 &\geq 11 \\
R_1 + 2R_2 + R_3 &\geq 11 \\
R_1 + R_2 + 2R_3 &\geq 11 \\
R_1 + R_2 + R_3 &\geq 9 \\
R_1 + R_2 + R_3 &\leq 11
\end{aligned}$$

By checking the inequalities with the condition of (4.7), the last of inequality is invalid. In fact the whole set of inequalities describing the sum are given by

$$\{(R_1, R_2, R_3) :$$

$$\begin{aligned}
R_i &\geq 1 \ \text{ for } \ 1 \leq i \leq 3 \\
R_1 + R_2 &\geq 4 \\
R_1 + R_3 &\geq 4 \\
R_2 + R_3 &\geq 4
\end{aligned}$$

$$\begin{aligned} 2R_1 + R_2 + R_3 &\geq 11 \\ R_1 + 2R_2 + R_3 &\geq 11 \\ R_1 + R_2 + 2R_3 &\geq 11 \\ R_1 + R_2 + R_3 &\geq 9 \qquad \} \end{aligned}$$

Note again that the first three ineqaulities above are not given by the convex hull of the polytope, but it is readily obtained from the summand sets $\mathbf{r^x}$.

Example 4.4

consider the addition of the three sets $\mathbf{r^x} + \mathbf{r^y} + \mathbf{r^z} = \mathbf{R} \in \mathbb{R}^4$, where $\mathbf{r^x} \in \mathbb{R}^4$ is

$$\{(r_1^x, r_2^x, r_3^x, r_4^x) : \quad \begin{aligned} r_1^x + r_2^x &\geq 2 \\ r_1^x + r_3^x &\geq 2 \\ r_2^x + r_4^x &\geq 2 \qquad \} \end{aligned}$$

$\mathbf{r^y} \in \mathbb{R}^4$ is

$$\{(r_1^y, r_2^y, r_3^y, r_4^y) : \quad \begin{aligned} r_1^y + r_2^y + r_4^y &\geq 3 \\ r_2^y + r_3^y + r_4^y &\geq 3 \\ r_1^y + r_2^y + r_4^y &\geq 3 \qquad \} \end{aligned}$$

$\mathbf{r^z} \in \mathbb{R}^4$ is,

$$\{(r_1^z, r_2^z, r_3^z, r_4^z) : r_i^z \geq 0 \text{ for } 1 \leq i \leq 4,\ r_1^z + r_2^z + r_3^z + r_4^z \geq 5\}$$

The corresponding extreme points for the sets are $\mathbf{E_x}, \mathbf{E_y}$ and $\mathbf{E_z}$ and they all have the same set of extreme direction $\mathbf{D}$ where

$$\begin{aligned} \mathbf{E_x} &= \{(0,2,2,0),(2,0,0,2),(2,2,0,0)\} \\ \mathbf{E_y} &= \{(0,0,3,3),(0,3,0,0),(3,0,0,3),(3,0,3,0),(1.5,0,1.5,1.5)\} \\ \mathbf{E_z} &= \{(5,0,0,0),(0,5,0,0),(0,0,5,0),(0,0,0,5)\} \\ \mathbf{D} &= \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\} \end{aligned}$$

The cardinality of the set of extreme point sum is huge (60) and is not list here.

By finding the convex hull of the sum of extreme points and express the half spaces in terms of inequalities, we have,

$$R_2 + R_4 \leq 10 \tag{4.9}$$
$$R_1 + R_2 + R_3 + R_4 \leq 15 \tag{4.10}$$
$$R_1 + R_2 + R_3 \geq 5 \tag{4.11}$$
$$2R_1 + 2R_2 + R_3 + R_3 \geq 17 \tag{4.12}$$
$$R_1 + 2R_2 + 2R_3 + R_4 \geq 15 \tag{4.13}$$
$$2R_1 + 3R_2 + 2R_3 + 2R_4 \geq 27 \tag{4.14}$$
$$R_1 + R_2 + R_3 + R_4 \geq 12 \tag{4.15}$$
$$R_1 + 2R_2 + R_4 \geq 7 \tag{4.16}$$
$$R_2 + R_3 + R_4 \geq 5 \tag{4.17}$$
$$R_1 + 2R_2 + R_3 + 2R_4 \geq 17 \tag{4.18}$$
$$R_1 + 2R_2 + R_3 + R_4 \leq 22 \tag{4.19}$$
$$R_2 + R_3 \leq 12 \tag{4.20}$$
$$R_1 + R_2 + R_4 \geq 5 \tag{4.21}$$
$$R_1 + R_2 \leq 12 \tag{4.22}$$

So the invalid inequalities are (4.9), (4.10), (4.19),(4.20) and (4.22). In fact the whole set of inequalities of the sum is given by

$$R_1 + R_2 \geq 2 \tag{4.23}$$
$$R_1 + R_3 \geq 2 \tag{4.24}$$
$$R_2 + R_4 \geq 2 \tag{4.25}$$
$$2R_1 + R_2 + R_3 \geq 7 \tag{4.26}$$
$$R_1 + R_2 + R_3 \geq 5$$
$$R_2 + R_3 + R_4 \geq 5$$
$$R_1 + R_2 + R_4 \geq 5$$

$$\begin{aligned}
R_1 + 2R_2 + R_4 &\geq 7 \\
R_1 + R_2 + R_3 + R_4 &\geq 12 \\
R_1 + 2R_2 + 2R_3 + R_4 &\geq 15 \\
R_1 + 2R_2 + R_3 + 2R_4 &\geq 17 \\
2R_1 + 2R_2 + R_3 + R_3 &\geq 17 \\
2R_1 + 3R_2 + 2R_3 + 2R_4 &\geq 27
\end{aligned}$$

Note that (4.23), (4.24), (4.25) and (4.26) is not given by the convex hull algorithm. They are the inequalities describing the set $\mathbf{r^x} + \mathbf{r^y}$.

In the above discussion we try to construct the sum from the summands. Now for a given sum, we want to find its possible summands. For a convex set $\mathbf{X} \subset \mathbb{R}^n$, we want to know is it possible to express each elements $\mathbf{x}$ of $\mathbf{X}$ in terms of the sum m of components $\mathbf{x_i} \in \mathbf{X_i}, i = 1, 2, \ldots, m$. If it is possible to do so, then $\mathbf{X} \subseteq \mathbf{X_1} + \mathbf{X_2} + \ldots + \mathbf{X_m}$. We will find that extreme points of the set also play an important role in the process.

Theorem 8 *Suppose* $\mathbf{X} = \mathbf{X_1} + \mathbf{X_2} + \ldots + \mathbf{X_m}$ *and* $\mathbf{X_i}, i = 1, 2, \ldots, m$ *are convex and all have the same set of extreme directions* $\{\mathbf{d_j}, j = 1, 2, \ldots, p\}$. *If* $\{\mathbf{x_i}, i = 1, 2, \ldots, q\}$ *is in* $\mathbf{X}$ *then*

$$\mathbf{x} = \sum_{i=1}^{q} \lambda_i \mathbf{x_i} + \sum_{j=1}^{p} \mu_j \mathbf{d_j}$$

where

$$\sum_{i=1}^{q} \lambda_i = 1,$$

$$\lambda_i \geq 0 \text{ for } 1 \leq i \leq q.$$

$$\mu_j \geq 0 \text{ for } 1 \leq j \leq p.$$

is also in $\mathbf{X}$.

Proof

$$\begin{aligned}
\mathbf{x} &= \sum_{i=1}^{q} \lambda_i \sum_{r=1}^{m} \mathbf{x_i^r} + \sum_{j=1}^{p} \mu_j \mathbf{d_j} \quad \text{where } \mathbf{x_i^r} \in \mathbf{X_r} \text{ for } 1 \leq r \leq m \\
&= \sum_{r=2}^{m} \sum_{i=1}^{q} \lambda_i \mathbf{x_i^r} + \left(\sum_{i=1}^{q} \lambda_i \mathbf{x_i^1} + \sum_{j=1}^{p} \mu_j \mathbf{d_j}\right)
\end{aligned}$$

$$= \sum_{r=2}^{m} \mathbf{x}'_\mathbf{r} + \mathbf{x}'_\mathbf{1} \quad \text{for some } \mathbf{x}'_\mathbf{r} \in \mathbf{X_r}, r = 1,2,\ldots,m\}$$
$$= \sum_{r=1}^{m} \mathbf{x}'_\mathbf{r}$$

The third equality in the above expression holds because $\mathbf{X_r}, r = 1, 2, \ldots, m$ are convex and if $\mathbf{x_i}$ is in $\mathbf{X_i}, \mathbf{x_i} + \sum_{j=1}^{l} \mu_j \mathbf{d_j}$ is also in $\mathbf{X_i}$ given $\mathbf{d_j}, j = 1, 2, \ldots, p$ are directions of the set.

With Theorem (8) and Theorem (5), it suffices to prove that a polyhedral set is a subset of the sum of m component polyhedral sets if its extreme points belongs to the sum. That is, if every extreme point of a polyhedral set can be expressed as the sum of m points each belongs to one of the m summands, then all points of the set can be as a sum of m points each belonging to one of the summands.

4.4 Algorithms to Enumerate Extreme Points and Decompose Tuples

As now we may appreciate the important role of the extreme points in the analysis of polyhedral set. Here we will introduce a simple algorithm to explicitly enumerate all the extreme points of the set.

Suppose an polyhedral set $\mathbf{X} \in \mathbb{R}^n$ is given by

$$\{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}, \text{where } \mathbf{A} \text{ is a } m \times n \text{ matrix} \tag{4.27}$$

That is $\mathbf{X}$ has m linear inequality constraints. In linear programming terms, this representation known as in *cannonical form.* An equivalent form of represenation is called *standard form* in which all the constraints are equations and all the variables are nonnegative. In the convertion from canonical form to standard form, a non-negative *surplus* variable s_i is added to an inequality of the form $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \geq b_i$ to convert it into an equation of form $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n - s_i = b_i$. It is easily seen that these two forms are equivalent since there is no extra constraint on s_i which is then allowed to assume any value grater than zero, so the correponding inequality in canonical form always holds. Standard form representation of $\mathbf{X}$ is

$$\{\mathbf{x} : \mathbf{A_s}\mathbf{x_s} = \mathbf{b}, \mathbf{x_s} \geq 0\},$$

where $\mathbf{A_s} = [\mathbf{A} \; -\mathbf{I}]$ is a $m \times (n+m)$ matrix, I is an identical matrix and $\mathbf{x_s} = [x_1 \; x_2 \; \ldots \; x_n \; s_1 \; s_2 \; \ldots \; s_m]$.

By possibly rearranging the columns of $\mathbf{A_s}$ and change it to $[\mathbf{B} \; \mathbf{N}]$ where $\mathbf{B}$ is an $m \times m$ invertible matrix. The solution $\mathbf{x} = [\mathbf{x_B} \; \mathbf{x_N}]^T$ to the equation $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b}^T$$

and

$$\mathbf{x_N} = 0$$

is called a *basic solution* of the system. If $\mathbf{x_B} \geq 0$ then $\mathbf{x}$ is called a *basic feasible solution* of the system. $\mathbf{B}$ is called a *basic matrix* of the system. The following theorem relates the basic feasible solution with extreme points of the systems. The proof can be found in [2].

Theorem 9 $\mathbf{x}$ *is an basic feasible solution of the system* $\mathbf{Ax{=}b}$ *if and only if* $\mathbf{x}$ *is an extreme point of the system.*

By forming all possible different basic matrices, we can find all the basic feasible solutions of the system.Since the system $\{\mathbf{x} : \mathbf{A_s x_s} = \mathbf{b}, \mathbf{x_s} \geq 0\}$ is equvalent to $\{\mathbf{x} : \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq 0\}$, the exterme points of the latter are obtained by picking $[x_1 \; x_2 \; \ldots \; x_n]$ from the basic feasible sulutions $\mathbf{x} = [\mathbf{x_B} \; \mathbf{x_N}]^T$. Since there are a maximun of C_m^{m+n} different basic matrix can be chosen from a $m \times (m+n)$ matrix,the number of extreme point of the system in (4.27) is less than or equal to C_m^{m+n}. The computational effort to exhaust all the extreme points can be large since the number of extreme points may grow exponentially with the dimension of the set n and the number of the inequalities of the system m. For the size of the problems we consider in this thesis, this algorithm is relatively primitive but yet very useful.

Now for a k-level-m-encoder MDCS with certain admissiblity condition, given an admissible coding rate m-tuple, one may want to know whether it is achievable by superposition. To implement superposition means we have to decompose the m-tuple into k parts, each satifying the admissiblity of one of the data streams, that is to allocate the the sufficient coding rate in each encoder for the admissibility of each data stream.

We want to find whether a m-tuple $\mathbf{t} = [t_1 \; t_2 \; \ldots \; t_m]$ is inside the convex set

$$\mathbf{X} = \mathbf{X_1} + \mathbf{X_2} + \ldots + \mathbf{X_k} \text{ where } \mathbf{X_i} \in \mathbb{R}^n, i = 1, 2, \ldots, k \text{ is polyhedral sets.}$$

If it is so, we may also want to know explicitly how to decompose $\mathbf{t}$ into k parts each belonging to one of $\mathbf{X_i}$. Now we set up a LP problem which is able to give answer to our question. Denote an element of $\mathbf{X_i} \in \mathbb{R}^m$ as $\mathbf{x^i} = [x_1^i \; x_2^i \; \ldots \; x_m^i]$ and the set is represented by the system $\{\mathbf{x^i} : \mathbf{B_i x^i} \geq \mathbf{b_i}\}$, consider the LP problem

$$\begin{aligned} min \;\; c &= \sum_{i=1}^{k}\sum_{j=1}^{m} x_j^i \\ \text{subject to} & \\ \sum_{i=1}^{k} x_j^i &= t_j \;\text{ for } 1 \leq j \leq m \\ \mathbf{B_i x^i} &\geq \mathbf{b_i} \;\text{ for } 1 \leq i \leq k \\ x_j^i &\geq 0 \end{aligned}$$

Now, by checking whether the feasible region is nonempty, we know whether the tuple $\mathbf{t} = [t_1 \; t_2 \; \ldots \; t_m]$ is inside set $\mathbf{X}$. Also by explicitly solve the LP problem using standard simplex method, we can obtain the set of values $\mathbf{x^i} \in \mathbb{R}^m$ such that $\mathbf{t} = \sum_{i=1}^{k} \mathbf{x^i}$. When setting up the LP problem, minimizing or maximizing the objective function is indifferent to our problem at the moment.

Here we discuss further how we can check that the feasible region is nonempty. By converting the constraints describing the feasible region of the above LP problem to the standard form, we can express the feasible region as

$$\mathbf{Ax} = \mathbf{b}$$

where

$\mathbf{A}$ is a $q \times (km + p)$ matrix and $\mathbf{x} = [x_1^1 \; x_2^1, \ldots \; x_m^1 \; x_1^2 \; \ldots \; x_m^2 \; \ldots \; x_m^k \; s_1 \; s_2 \; \ldots \; s_p]$

and p is the number of inequality constraints in the original LP problem. q is the total number of equality and inequality constraints. $km + p$ is the total number of variables in the standard form representation. Now we state without proof a famous lemma in Linear Programming known as the Farkas' Lemma. There are a few variants of it and we just state the one suitable for solving the problem here.

Lemma 1 *Assume* $\mathbf{A}$ *is a* $m \times n$ *matrix.*

Either $\{\mathbf{x} \in \mathbb{R}^{+n} : \; \mathbf{Ax} = \mathbf{b}\} \neq \phi$ *or* $\{\mathbf{y} \in \mathbb{R}^m : \; \mathbf{yA} \geq \mathbf{0}, \mathbf{yb} < \mathbf{0}\} \neq \phi$.

If the solution of the LP problem $\{\min \mathbf{c} = \mathbf{yb}\ s.t.\ \mathbf{yA} \geq \mathbf{0}\}$ is zero which must occur at $\mathbf{y} = \mathbf{0}$ since $\mathbf{b} > \mathbf{0}$, we can know that the region $\{\mathbf{y} \in \mathbb{R}^m : \mathbf{yA} \geq \mathbf{0}, \mathbf{yb} < \mathbf{0}\}$ is empty which in turn implies that the region $\{\mathbf{x} \in \mathbb{R}^{+n} : \mathbf{Ax} = \mathbf{b}\}$ is nonempty. So by applying standard optimality test to the LP problem $\{\min \mathbf{c} = \mathbf{yb}\ s.t.\ \mathbf{yA} \geq \mathbf{0}\}$ at the point $\mathbf{y} = \mathbf{0}$, we can know whether the feasible region is nonempty.

Chapter 5

Conclusion and Further Research

5.1 Conclusion

In this thesis, multilevel diversity coding systems in which the information source consist of independent data streams are studied. We show noiseless coding of the different data streams by *superposition*, which is generally feasible in all MDCs, is also optimal in most of the cases we studied. We looked into detail the two equivalent representation of the coding rate region by subrate constraints and rate constraints. A crucial step in the proof of the converse of the coding scheme by superposition is to establish the rate constraints from the subrate constraints which are evident once the MDCS is defined. We introduced mathematical tools from convex set analysis and linear programming to solve the problem and so enabled us to prove that superposition is the optimal scheme in all the positive examples given in the thesis. We also look into some examples of another MDCS coding scheme, *linear combination* of data streams. A class of MDCS's where superposition is always not optimal are discovered and the optimality can be achieved by linear combination. We focus on the 3-encoder MDCS's and discover that superposition and linear combination are two complementary optimal coding scheme in 2-level-3-encoder MDCS's. Also, a coding scheme applying both linear combination and superpostion also complements superpostion as the optimal coding schemes in all 3-level-3-encoder MDCS's. We are able to characterize the addmissible coding rate region for all the cases above. We are unable to derive the general necessary and sufficient condition for the optimality of superposition but hope that the work done at the moment can shed light for further investigation.

In addition, we look into a special class of MDCS's in which the encoders and decoders have symmetrical connectivity (SMDCS). This class of MDCS's have special applications and the rate constraints describing the coding rate region also has symmetrical structure. We divide the SMDCS into different subclasses and the admissible coding rate regions are found for almost all the subclasses up to 4 levels of decoders. Coding by the principle of superposition was found to be optimal in all the SMDCS cases we have studied in the thesis.

In the analysis of the coding rate region which are in fact some polyhedral sets with special properties. We propose the characterization of the rate region by extreme points and extreme directions. The addition operation of polyhedral sets was analyzed which is crucial in analyze the coding rate region induced by superposition. Some basic theorems are stated and some are proved and primitive algorithms presented to find the extrme points of the sets and check the feasibility of superposition for a rate tuple.

5.2 Suggestions for Further Research

1. The necessary and sufficient condition for the optimality of superposition is yet to be found given that we have discover more and more examples that it is optimal. Another coding scheme in MDCS's is linear combination of data streams which can be applied when superposition is not optimal in the examples studied in the thesis. We cannot generalize the result to the k-level-m-encoder MDCS's. This problem is intrinsically difficult since not all MDCS can apply linear combination and to see whether superposition is optimal involves some nontrival processes. Even superposition is not optimal, it is also not easy to find an optimal substitute. The language used to describe the feasiblity of linear combination in MDCS (one example of which is sequentially refinable MDCS) and the optimality of superposition as studied in this thesis is not similar at all. The characterization of sequentially refinable MDCS involve connection structure of encoders and decoders [12] while characterization of superposition in MDCS involves the description of the coding rate region in terms of subrate or rate constraints. And how does connection structure relates to the coding rate region description is still unclear. So we are unable to solve this open problem

at the moment. In this thesis we prove that a class of MDCS's where linear combination can apply and superposition is always not optimal, in addtion to a large amount of examples, which may help to solve the problem in the future.

2. To establish the rate constraints of the coding rate region is a vital step in the proof of converse of the coding scheme. This invloves addition of sets of described by different subrate constraints. The charecterization of the coding rate region by extreme points and extreme direction is a feasible solution to this problem. Fast algorithms to enumerate the extreme points are needed. Direct enumeration of the extreme points of the sets without relying on computer is even better.

3. We have explored the class of symmetrical MDCS denoted by

$$SMDCS[k, m, (m_1, m_2, \ldots, m_k)]$$

   We are able solve a few subclasses of SMDCS but are unable to solve the general probelm. We discover that in this class the connectivity of the encoders and decoders, the rate constraints and the structure of the extreme points all have special symmetrical structure and some correlation between the three are discovered. The general problem may be solved once the relation between the three is discovered. But it is again a very difficult problem.

Appendix A

Proof of Equivalence of $\mathbf{r_{sp}}$ and $\mathbf{R_{sp}}$ in Chapter 3

Here we prove that $\mathbf{r_{2m1m}}$, $\mathbf{r_{33123}}$ and $\mathbf{r_{441234}}$ are in fact equivalent to $\mathbf{R_{2m1m}}$, $\mathbf{R_{33123}}$, and $\mathbf{R_{441234}}$ respectively as stated in chapter 3. In fact there are many $\mathbf{r_{sp}}$ in this thesis which are not so obviously seen to be equivalent to their coresponding $\mathbf{R_{sp}}$ but we are unable to prove their equivalence one by one since it will be to lengthy to do so. We choose three of them to illustrate the process of proof. The proof involve two main steps. First we prove that $\mathbf{r_{sp}}$ is a subset of $\mathbf{R_{sp}}$ by showing that the inequalities of $\mathbf{R_{sp}}$ is implied by the inequalities in $\mathbf{r_{sp}}$. Also from the discussion in Chapter 4, we can see that all sets of the same dimension have the same set of extreme directions since they are in positive region $\mathbf{R^{n+}}$ and the coeficients of all the inequalities (with $\geq$ sign) are positive. Given these conditions we can invoke Theorem (8) and prove that $\mathbf{R_{sp}}$ is a subset of $\mathbf{r_{sp}}$ by enumerate all the extreme points of $\mathbf{R_{sp}}$ and show that they are in $\mathbf{r_{sp}}$.

A.1 $\mathbf{r_{2m1m}}$ and $\mathbf{R_{2m1m}}$

$\mathbf{r_{2m1m}}$ is given by

$\{(R_1, R_2, \ldots, R_m):$

$$r_i^x \geq H(X) \quad \text{for } 1 \leq i \leq m \tag{A.1}$$

$$r_1^y + r_2^y + \ldots + r_m^y \geq H(Y) \quad \} \tag{A.2}$$

and $\mathbf{R_{2m1m}}$ is given by

$\{(R_1, R_2, \ldots, R_m):$

$$R_i \geq H(X) \quad \text{for } 1 \leq i \leq m \tag{A.3}$$

$$R_1 + R_2 + \ldots + R_m \geq kH(X) + H(Y) \quad \} \tag{A.4}$$

It is easily seen that (A.1) implies (A.3) and $m \times (A.1) + (A.2)$ implies (A.4) given that $R_i = r_i^x + r_i^y$ for $1 \leq i \leq m$. So $\mathbf{r_{2m1m}} \subseteq \mathbf{R_{2m1m}}$.

Now there are m extreme points in $\mathbf{R_{2m1m}}$ and we denote each as Q_i. They are

$$\begin{aligned}
Q_1 &= (H(X)+H(Y),\ H(X),\ \ldots,\ H(X)) & &= (H(X),\ H(X),\ \ldots,\ H(X)) \\
& & &\quad +(H(Y),\ 0,\ \ldots,\ 0) \\
Q_2 &= (H(X),\ H(X)+H(Y),\ H(X),\ \ldots,\ H(X)) & &= (H(X),\ H(X),\ \ldots,\ H(X)) \\
& & &\quad +(0,\ H(Y),\ 0,\ \ldots,\ 0) \\
&\vdots & &\vdots \\
Q_m &= (H(X),\ \ldots,\ H(X),\ H(X)+H(Y)) & &= (H(X),\ H(X),\ \ldots,\ H(X)) \\
& & &\quad +(0,\ \ldots,\ 0,\ H(Y))
\end{aligned}$$

So all extreme points are in $\mathbf{r_{2m1m}}$. So $\mathbf{R_{2m1m}} \subseteq \mathbf{R_{2m1m}}$. The two sets are equivalent.

A.2 $\mathbf{r_{33123}}$ and $\mathbf{R_{33123}}$

$\mathbf{r_{33123}}$ is given by

$\{(R_1, R_2, R_3):\ R_i = r_i^x + r_i^y + r_i^z$ where $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1, 2, 3$ and

$$r_1^x \geq H(X) \tag{A.5}$$

$$r_2^x \geq H(X) \tag{A.6}$$

$$r_3^x \geq H(X) \tag{A.7}$$

$$r_1^y + r_2^y \geq H(Y) \tag{A.8}$$

$$r_2^y + r_3^3 \geq H(Y) \tag{A.9}$$

$$r_1^x + r_3^x \geq H(Y) \tag{A.10}$$

$$r_1^z + r_2^z + r_3^z \geq H(Z) \quad \} \tag{A.11}$$

$\mathbf{R_{33123}}$ is given by

$$
\begin{aligned}
\{(R_1, R_2, R_3) : R_i \geq 0 \text{ for } i = 1,2,3 \quad \text{and} \\
R_1 &\geq H(X) && \text{(A.12)} \\
R_2 &\geq H(X) && \text{(A.13)} \\
R_3 &\geq H(X) && \text{(A.14)} \\
R_1 + R_2 &\geq 2H(X) + H(Y) && \text{(A.15)} \\
R_2 + R_3 &\geq 2H(X) + H(Y) && \text{(A.16)} \\
R_1 + R_3 &\geq 2H(X) + H(Y) && \text{(A.17)} \\
R_1 + R_2 + R_3 &\geq 3H(X) + \frac{3}{2}H(Y) + H(Z) && \text{(A.18)} \\
2R_1 + R_2 + R_3 &\geq 4H(X) + 2H(Y) + H(Z) && \text{(A.19)} \\
R_1 + 2R_2 + R_3 &\geq 4H(X) + 2H(Y) + H(Z) && \text{(A.20)} \\
R_1 + R_2 + 2R_3 &\geq 4H(X) + 2H(Y) + H(Z) \quad \} && \text{(A.21)}
\end{aligned}
$$

We will first prove $\mathbf{r_{33123}} \subseteq \mathbf{R_{33123}}$ by showing that (A.5)-(A.11) imply (A.12)-(A.21) given that there exist $r_i^x, r_i^y, r_i^z \geq 0$ for $i = 1,2,3$ such that $R_i = r_i^x + r_i^y + r_i^z$ is satisfied.

Obviously, $(A.5) \Rightarrow (A.12), (A.6) \Rightarrow (A.13), (A.7) \Rightarrow (A.14)$

Also,

$$
\begin{aligned}
(A.5) + (A.6) + (A.8) &\Rightarrow (A.15) \\
(A.5) + (A.7) + (A.9) &\Rightarrow (A.16) \\
(A.6) + (A.7) + (A.10) &\Rightarrow (A.17) \\
2 \times (A.5) + (A.6) + (A.7) + (A.8) + (A.9) + (A.11) &\Rightarrow (A.19) \\
(A.5) + 2 \times (A.6) + (A.7) + (A.8) + (A.10) + (A.11) &\Rightarrow (A.20) \\
(A.5) + (A.6) + 2 \times (A.7) + (A.9) + (A.10) + (A.11) &\Rightarrow (A.21) \\
(A.5) + (A.6) + (A.7) + \tfrac{1}{2} \times (A.8) + \tfrac{1}{2} \times (A.9) + \tfrac{1}{2} \times (A.10) + (A.11) &\Rightarrow (A.18)
\end{aligned}
$$

In the following, we list all the extreme points (cf. Figure 3.4.) of $\mathbf{R_{33123}}$ and show that they are in fact in $\mathbf{r_{33123}}$:

$$
\begin{aligned}
Q_1 &= (H(X),\ H(X)+H(Y),\ H(X)+H(Y)+H(Z)) \\
&= (H(X),\ H(X),\ H(X)) + (0,\ H(Y),\ H(Y)) + (0,\ 0,\ H(Z)) \\
Q_2 &= (H(X),\ H(X)+H(Y)+H(Z),\ H(X)+H(Y)) \\
&= (H(X),\ H(X),\ H(X)) + (0,\ H(Y),\ H(Y)) + (0,\ H(Z),\ 0) \\
Q_3 &= (H(X)+H(Y),\ H(X)+H(Y)+H(Z),\ H(X)) \\
&= (H(X),\ H(X),\ H(X)) + (H(Y),\ 0,\ H(Y)) + (0,\ 0,\ H(Z)) \\
Q_4 &= (H(X)+H(Y)+H(Z),\ H(X),\ H(X)+H(Y)) \\
&= (H(X),\ H(X),\ H(X)) + (H(Y),\ 0,\ H(Y)) + (H(Z),\ 0,\ 0) \\
Q_5 &= (H(X)+H(Y),\ H(X)+H(Y)+H(Z),\ H(X)) \\
&= (H(X),\ H(X),\ H(X)) + (H(Y),\ H(Y),\ 0) + (0,\ H(Z),\ 0) \\
Q_6 &= (H(X)+H(Y)+H(Z),\ H(X)+H(Y),\ H(X)) \\
&= (H(X),\ H(X),\ H(X)) + (H(Y),\ H(Y),\ 0) + (H(Z),\ 0,\ 0) \\
Q_7 &= (H(X)+\tfrac{H(Y)}{2},\ H(X)+\tfrac{H(Y)}{2},\ H(X)+\tfrac{H(Y)}{2}+H(Z)) \\
&= (H(X),\ H(X),\ H(X)) + (\tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2}) + (0,\ 0,\ H(Z)) \\
Q_8 &= (H(X)+\tfrac{H(X)}{2},\ H(X)+\tfrac{H(Y)}{2}+H(Z),\ H(X)+\tfrac{H(Y)}{2}) \\
&= (H(X),\ H(X),\ H(X)) + (\tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2}) + (0,\ H(Z),\ 0) \\
Q_9 &= (H(X)+\tfrac{H(Y)}{2}+H(Z),\ H(X)+\tfrac{H(Y)}{2},\ H(X)+\tfrac{H(Y)}{2}) \\
&= (H(X),\ H(X),\ H(X)) + (\tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2}) + (H(Z),\ 0,\ 0)
\end{aligned}
$$

Therefore $\mathbf{R_{33123}} \subseteq \mathbf{r_{33123}}$.

A.3 $\mathbf{r_{441234}}$ and $\mathbf{R_{441234}}$

$\mathbf{r_{441234}}$ is given by
$\{(R_1, R_2, R_3, R_4):$

$$R_i = r_i^w + r_i^x + r_i^y + r_i^z \quad \text{for } 1 \le i \le 4 \tag{A.22}$$

and

$$r_i^w \ge H(W) \quad \text{for } 1 \le i \le 4 \tag{A.23}$$
$$r_i^x + r_j^x \ge H(X) \quad \text{for } 1 \le i < j \le 4 \tag{A.24}$$
$$r_i^y + r_j^y + r_k^y \ge H(Y) \quad \text{for } 1 \le i < j < k \le 4 \tag{A.25}$$
$$r_1^z + r_2^z + r_3^z + r_4^z \ge H(Z) \quad \} \tag{A.26}$$

Define $\pi(1234)$ as the set of all permutation of the sequence (1,2,3,4). The coding rate region described by rate sum inequalities, denoted as $\mathbf{R_{441234}}$, is given by

$\{(R_1, R_2, R_3, R_4):$

$$
\begin{aligned}
R_i &\geq H(W) \quad \text{for } 1 \leq i \leq 4 && \text{(A.27)}\\
R_i + R_j &\geq 2H(W) + H(X) \quad \text{for } 1 \leq i < j \leq 4 && \text{(A.28)}\\
R_i + R_j + R_k &\geq 3H(W) + \frac{3}{2}H(X) + H(Y) && \text{(A.29)}\\
2R_i + R_j + R_k &\geq 4H(X) + 2H(Y) + H(Z) && \text{(A.30)}\\
R_i + 2R_j + R_k &\geq 4H(X) + 2H(Y) + H(Z) && \text{(A.31)}\\
R_i + R_j + 2R_k &\geq 4H(X) + 2H(Y) + H(Z) && \text{(A.32)}\\
R_1 + R_2 + R_3 + R_4 &\geq 4H(W) + 2H(X) + \frac{4}{3}H(Y) + H(Z) && \text{(A.33)}\\
2R_i + 2R_j + R_k + R_l &\geq 6H(W) + 3H(X) + 2H(Y) + H(Z) && \text{(A.34)}\\
3R_i + R_j + R_k + R_l &\geq 6H(W) + 3H(X) + \frac{3}{2} + H(Z) && \text{(A.35)}\\
3R_i + 2R_j + 2R_k + 2R_l &\geq 7H(W) + \frac{9}{2}H(X) + 3H(Y) + 2H(Z) && \text{(A.36)}\\
4R_i + 2R_j + R_k + R_l &\geq 8H(W) + 4H(X) + 2H(Y) + H(Z) && \text{(A.37)}\\
&\quad \forall (i,j,k,l) \in \pi(1,2,3,4) \qquad \}
\end{aligned}
$$

We will first prove $\mathbf{r_{441234}} \subseteq \mathbf{R_{441234}}$ by showing that (A.24)-(A.26) imply (A.27)-(A.37) given that there exist $r_i^w, r_i^x, r_i^y, r_i^z \geq 0$ for $1 \leq i \leq 4$ such that (A.22) is satisfied.

Obviously,

$$
\begin{aligned}
(A.23) &\Rightarrow (A.27)
\end{aligned}
$$

Also,

$$
\begin{aligned}
2 \times (A.23) + (A.24) &\Rightarrow (A.28)\\
3 \times (A.23) + \tfrac{3}{2} \times (A.24) + (A.25) &\Rightarrow (A.29)\\
4 \times (A.23) + 2 \times (A.24) + (A.25) &\Rightarrow (A.30), (A.31), (A.32)\\
4 \times (A.23) + 2 \times (A.24) + \tfrac{4}{3} \times (A.25) + (A.26) &\Rightarrow (A.18)\\
6 \times (A.23) + 3 \times (A.24) + 2 \times (A.25) + (A.26) &\Rightarrow (A.34)\\
6 \times (A.23) + 3 \times (A.24) + \tfrac{3}{2} \times (A.25) + (A.26) &\Rightarrow (A.35)\\
7 \times (A.23) + \tfrac{9}{2} \times (A.24) + 3 \times (A.25) + 2 \times (A.26) &\Rightarrow (A.36)\\
8 \times (A.23) + 4 \times (A.24) + 2 \times (A.25) + (A.26) &\Rightarrow (A.37)
\end{aligned}
$$

Now we list all the extreme points of $\mathbf{R_{441234}}$, and show that they are in $\mathbf{r_{441234}}$. It is interesting to notice that like the inequalities describing the set, the extreme points also have symmetrical structures. We list one extreme points in each different typical symmetrical pattern and a $(\times n)$ is put in the end of each listed point to indicate that there are n similar points with the same pattern. Note that there are totally 64 extreme points in $\mathbf{R_{441234}}$.

$$
\begin{aligned}
Q_1 &= (H(W),\ H(W)+H(X),\ H(W)+H(X)+H(Y),\ H(W)+H(X)+H(Y)+H(Z))\ \ (\times 24)\\
&= (H(W),\ H(W),\ H(W),\ H(W))+(0,\ H(X),\ H(X),\ H(X))\\
&\quad +(0,\ 0,\ H(Y),\ H(Y))+(0,\ 0,\ 0,\ H(Z))\\
Q_2 &= (H(W),\ H(W)+H(X)+\tfrac{H(Y)}{2},\ H(W)+H(X)+\tfrac{H(Y)}{2},\ H(W)+H(X)+\tfrac{H(Y)}{2}+H(Z))\ \ (\times 12)\\
&= (H(W),\ H(W),\ H(W),\ H(W))+(H(X),\ H(X),\ H(X))\\
&\quad +(0,\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2})+(0,\ 0,\ 0,\ H(Z))\\
Q_3 &= (H(W)+\tfrac{H(X)}{2},\ H(W)+\tfrac{H(X)}{2},\ H(W)+\tfrac{H(X)}{2}+H(Y),\ H(W)+\tfrac{H(X)}{2}+H(Y)+H(Z))\ \ (\times 12)\\
&= (H(W),\ H(W),\ H(W),\ H(W))+(\tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2})\\
&\quad +(0,\ 0,\ H(Y),\ H(Y))+(0,\ 0,\ 0,\ H(Z))\\
Q_4 &= (H(W)+\tfrac{H(X)}{2},\ H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{2},\ H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{2},\ H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{2}+H(Z))\ \ (\times 12)\\
&= (H(W),\ H(W),\ H(W),\ H(W))+(\tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2})\\
&\quad +(0,\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2},\ \tfrac{H(Y)}{2})+(0,\ 0,\ 0,\ H(Z))\\
Q_5 &= (H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{3}, H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{3}, H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{3}, H(W)+\tfrac{H(X)}{2}+\tfrac{H(Y)}{3}+H(Z))(\times\\
&= (H(W),\ H(W),\ H(W),\ H(W))+(\tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2},\ \tfrac{H(X)}{2})\\
&\quad +(\tfrac{H(Y)}{3},\ \tfrac{H(Y)}{3},\ \tfrac{H(Y)}{3},\ \tfrac{H(Y)}{3})+(0,\ 0,\ 0,\ H(Z))
\end{aligned}
$$

Appendix B

A Class of MDCS Where Superposition is Always Not Optimal

We have seen in Chapter 1 and Chapter 2 that superposition is not optimal in some 3-encoder-MDCS's. But at the moment we cannot find the general necesssary and sufficient condition for the optimality of superposition. In 2-level-3-encoder MDCS's, we discover two configurations in which superposition is not optimal but the optimality is achieved by linear combination. We discover that these two cases are members of a more general class of MDCS where linear combination is optimal while superposition is always not optimal. (Note the subtlety is that the optimality of linear combination and superposition in general is not mutually conflicting, see Section 1.7.) Regonition of this class of MDCS provide us with further insight on the necessary conditions of the optimality of superposition which remains an open problem.

In such class of MDCS, the information source is $\{(X_k^1, X_k^2, \ldots, X_k^m)\}$ where $\{X_k^i\}$ $i = 1, 2, \ldots m$ are independent data streams with the same entropy rate. There are at least $m+1$ encoders and $m+2$ decoders in the MDCS. The basic structure of the MDCS are described in the following.

Decoder i is connected to encoder 1 to encoder i, and recovers $\{(X_k^1, X_k^2, \ldots, X_k^i)\}$ for $1 \leq i \leq m$. Decoder $m+1$ are connected to encoder $1, 2, \ldots, m-1, m+1$ and recovers $\{(X_k^1, X_k^2, \ldots, X_k^m)\}$. Moreover decoder $m+2$ are connected to encoder $2, 3, \ldots, m+1$ and recovers $\{(X_k^1, X_k^2, \ldots X_k^i)\}$ where i can by any number in $\{1, 2, 3, \ldots, m\}$.

Assume the entropy rate of each data stream is H. The coding rate region $\mathbf{r_{sp}}$ is given by

$\{(R_1, R_2, \ldots, R_m): \; R_i = r_i^1 + r_i^2 + \ldots + r_i^m$
where
$r_i^j \geq 0$ for $i = 1, 2, \ldots, m+2, \; j = 1, 2, \ldots, m$ and

$$r_1^1 \geq H \tag{B.1}$$
$$r_1^2 + r_2^2 \geq H \tag{B.2}$$
$$r_1^3 + r_2^3 + r_3^3 \geq H \tag{B.3}$$
$$\vdots$$
$$r_1^m + r_2^m + \ldots + r_m^m \geq H \tag{B.4}$$
$$r_1^m + r_2^m + \ldots + r_{m-1}^m + r_{m+1}^m \geq H \tag{B.5}$$
$$r_2^1 + r_3^1 + \ldots + r_{m+1}^1 \geq H \tag{B.6}$$
$$\vdots$$
$$r_2^i + r_3^i + \ldots + r_{m+1}^i \geq H \quad \} \tag{B.7}$$

The rate tuple $(H, H, \ldots, H)$ is admissible by coding the data streams in encoder $1, 2, \ldots, m+1$ as $\{X_k^1\}, \{X_k^2\}, \ldots, \{X_k^m\}$ and $\{X_k^1 \oplus X_k^2 \oplus \ldots \oplus X_k^m\}$ respectively but this rate tuple is not in $\mathbf{r}_{\mathrm{sp}}$. We will show this in the following.

With the constraint

$$R_i = r_i^1 + r_i^2 + \ldots + r_i^m = H \tag{B.8}$$

for $i = 1$ and (B.1) we must have $r_1^1 = H$ and

$$r_1^2 = r_1^3 = \ldots = r_1^m = 0$$

With $r_1^2 = 0$ and (B.2) we must have $r_2^2 = H$ and by (B.8) for $i = 2$ we must have

$$r_2^1 = r_2^3 = \ldots = r_2^m = 0$$

With $r_1^3 = r_2^3 = 0$ and (B.3) we must have $r_3^3 = H$ and by (B.8) for $i = 3$ we must have

$$r_3^1 = r_3^2 = r_3^4 = \ldots = r_3^m = 0$$

Applying similar arguments consecutively we finally have

$$r_1^1 = r_2^2 = r_3^3 = \ldots = r_m^m = H$$

$$r_2^1 = r_3^1 = r_4^1 = \ldots = r_m^1 = H \tag{B.9}$$

and

$$r_1^m = r_2^m = \ldots = r_{m-1}^m = 0$$

with this and (B.5) we must have

$$r_{m+1}^m = H$$

With this and (B.8) for $i = m + 1$ we have

$$r_{m+1}^1 = r_{m+1}^2 = \ldots = r_{m+1}^{m-1} = 0$$

With this and (B.9) we have

$$r_2^1 = r_3^1 = \ldots = r_{m+1}^1 = 0$$

But in this case (B.6) is not satisfied. Thus $(H, H, \ldots, H)$ is not in $\mathbf{r_{sp}}$. So superposition is not optimal in this class of MDCS. Note that constraint (B.7) is not used for $i = 2, 3, \ldots, m$ so the value of i is indifferent as long as i is greater than or equal to 1.

The two examples of 2-level-3-encoder MDCS's for which superpositions is not optimal in Chapter 2 are obtained by setting $m = 2$, $(X^1 = X, X^2 = Y)$ and $i = 1, 2$ (Example 2.1.5) and $i = 1$ for (Example 2.1.6) respectively.

By the arguments in section 2 of chapter 2, superposition is also not optimal in an MDCS with any embedded MDCS having the same configuration as any member in this class of MDCS. So in fact we can construct many examples of MDCS's where superposition is not optimal.

Bibliography

[1] E. Ayanoglu, R. D. Gitlin, C.-L. I, and J. Mazo, "Diversity coding for transparent self-healing and fault-tolerant communication networks," IEEE International Symposium on Information Theory, San Diego, CA, Jan 1990.

[2] Mokhtar S. Bazaraa, John J. Jarvis, Hanif D. Sherali, *Linear Programming and Network Flows.* Wiley, 1990.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, New York, 1991.

[4] A. J. Goldman, "Resolution and separation theorems for polyhedral convex sets," in *Linear Inequalities and Related Systems*, H. W. Kuhn and A. W. Tucker (Eds), Princeton University Press, Princeton, N. J., 1956.

[5] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Inform. Contr.*, vol. 36, pp. 133-156, 1978.

[6] Paul J. Kelly, Max L. Weiss, *Geometry and Convexity* .Wiley, 1979.

[7] T. T. Lee and S. C. Liew, "Parallel communications for ATM network control and management," GLOBECOM '93, Houston, TX, Nov 29-Dec 2, 1993.

[8] E.D. Karnin, J.Greene, and M.E. Hellman, "On secret sharing system," *IEEE Trans. Inform. Theory*, vol. 29, pp.35-41,1983.

[9] Franco P.Preparata, Micheal Ian Shamos, *Computational Geometry.* Springer-Verlag, New York, 1985.

[10] M. O. Rabin, "Efficient dispersal of information for security, load balancing and fault-tolerance," *J. ACM*, vol. 36, pp. 335-348, 1989.

[11] J. R. Roche, A. Dembo, and A. Nobel, "Distributed information storage," IEEE International Symposium on Information Theory, Kobe, Japan, June 1988.

[12] J. R. Roche, "Distributed information storage," Ph.D. thesis, Stanford University, March 1992.

[13] J. R. Roche, *private communication*, 1993.

[14] James R. Roche, R. W. Yeung and Ka Pun Hau "Symmetrical multilevel diversity coding ", presented in IEEE Information Theory Workshop June 25-29 1995. Rydzyna, Poland.

[15] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, pp. 612-613, 1979.

[16] C. E. Shannon, "A mathematical theory of communication," Bell Sys. Tech. Journal, vol.27, pp.379-423, 623-656,1948.

[17] R. C. Singleton, "Maximum distance q-nary codes," *IEEE Trans. Inform. Theory*, vol. 10, pp. 116-118, 1964.

[18] R. W. Yeung, "Multilevel diversity coding with distortion", *IEEE Trans. Inform. Theory*, vol. 41, pp.412-422, 1995.